

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:00:40 ; Search time 34.94 Seconds  
(without alignments)  
47.963 Million cell updates/sec

Title: US-09-763-397A-25

Perfect score: 108

Sequence: 1 MKFLVNVLFVFMVVIYSIYAD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	94.4	70	1 MPB1	melittin, major, p
2	52	48.1	209	2 T40841	immunodominant pro
3	51	47.2	227	2 H69433	hypothetical prote
4	50.5	46.8	352	2 H66742	hypothetical prote
5	49	45.4	133	2 D70105	hypothetical prote
6	49	45.4	305	1 NKVLHH	core antigen - her
7	48	44.4	40	2 S44136	NADH dehydrogenase
8	48	44.4	88	2 S66287	sapacin B precurs
9	48	44.4	117	2 S82703	NADH dehydrogenase
10	47	43.5	127	1 VXECS	preprotein translo
11	47	43.5	127	2 E86089	preprotein translo
12	47	43.5	213	2 G69513	conserved hypothet
13	45	41.7	320	2 T26259	hypothetical prote
14	45	41.7	331	2 T28208	hypothetical prote
15	45	41.7	425	2 T15959	hypothetical prote
16	45	41.7	506	2 T23576	hypothetical prote
17	45	41.7	1078	2 A56715	calcium receptor (
18	45	41.7	1079	2 I59362	calcium/polyvalent
19	45	41.7	1085	2 S40476	Ca(2+)-sensing rec
20	45	41.7	1088	2 B56715	calcium receptor (
21	44	40.7	218	2 B69934	conserved hypothet
22	44	40.7	295	2 C54134	sapc protein homol
23	44	40.7	373	2 T16529	hypothetical prote
24	44	40.7	492	2 T23523	hypothetical prote
25	44	40.7	580	2 T28725	hypothetical prote
26	43.5	40.3	576	2 T41407	membrane transport
27	43.5	40.3	580	2 T40296	hypothetical prote
28	43	39.8	217	2 T19005	membrane transport
29	43	39.8	269	2 T10253	membrane protein M

ALIGNMENTS

RESULT 1

MPB1

melittin, major, precursor [validated] - honeybee

N:Contains: melittin F

C:Species: Apis mellifera (honeybee)

C>Date: 19-Feb-1984 #sequence\_revision 19-Feb-1984 #text\_change 15-Sep-2000

C:Accession: A91133; A91640; B61285; S23131; A01761; A18880

R:Vlasak, R.; Unger-Ullmann, C.; Kreil, G.; Frischauf, A.M.

Eur. J. Biochem. 135, 123-126, 1983

A>Title: Nucleotide sequence of cloned cDNA coding for honeybee prepro-melittin.

A:Reference number: A91133; MUID:83287387

A:Accession: A91133

A:Molecule type: mRNA

A:Residues: 1-70 <VLA>

A:Cross-references: GB:X02007; NID:95621; PIDN:CAA26038.1; PID:95622

R:Habermann, E.; Jentsch, J.

Hoppe-Seyler's Z. Physiol. Chem. 348, 37-50, 1967

A>Title: Sequenzanalyse des Melittins aus den tryptischen und peptischen Spaltstuecke

A:Reference number: A91640; MUID:68327913

A:Accession: A91640

A:Molecule type: protein

A:Residues: 44-69 <HAB>

R:Gauldie, J.; Hanson, J.M.; Shipolini, R.A.; Vernon, C.A.

Eur. J. Biochem. 83, 405-410, 1978

A>Title: The structures of some peptides from bee venom.

A:Reference number: A91253; MUID:78126868

A:Accession: B61285

A:Molecule type: protein

A:Residues: 51-69 <GAU>

R:Ramalingam, K.; Bello, J.

Biochem. J. 284, 663-665, 1992

A>Title: Effect of permethylation on the haemolytic activity of melittin.

A:Reference number: S23131; MUID:92321983

A:Accession: S23131

A>Status: preliminary

A:Molecule type: protein

A:Residues: 44-69 <RAM>

R:Schroeder, E.; Luebk, K.; Lehmann, M.; Beetz, I.

Experientia 27, 764-765, 1971

A>Title: Haemolytic activity and action on the surface tension of aqueous solutions o

A:Reference number: A91266; MUID:72098668

A:Contents: annotation; synthesis

A>Note: the structure of melittin was confirmed by synthesis of a peptide with full h

R:Krell, G.; Kreil-Kiss, G.

Biochem. Biophys. Res. Commun. 27, 275-280, 1967

A>Title: The isolation of N-formylglycine from a polypeptide present in bee venom.

A:Reference number: A90165; MUID:67248282

A:Contents: annotation

A>Note: about 10% of melittin is formylated at the amino end

R:Luebk, K.; Matthes, S.; Kloss, G.

Experientia 27, 765-767, 1971

A>Title: Isolation and structure of N(alpha)-formyl melittin.

A:Reference number: A91267; MUID:72098669  
A:Contents: annotation; synthesis  
A:Note: N-formyl-melittin major was isolated from the venom and its structure was confirmed by X-ray crystallography, 2.0 angstroms, residues 44-69  
R:Mollay, C.; Vilas, U.; Krell, G. 1982  
Proc. Natl. Acad. Sci. U.S.A. 79, 2260-2263, 1982  
A:Title: Cleavage of honeybee prepro-melittin by an endoprotease from rat liver microsomes  
A:Reference number: A18880; MUID:82247982  
A:Contents: annotation; intact signal sequence after cleavage by partially purified signal peptidase  
R:Eisenberg, D.; Gribskov, M.; Terwilliger, T.C.  
submitted to the Brookhaven Protein Data Bank, October 1990  
A:Reference number: A50496; PDB:2MLT  
A:Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 44-69  
R:Terwilliger, T.C.; Eisenberg, D.  
J. Biol. Chem. 257, 6010-6015, 1982

A:Title: The structure of melittin: I. Secondary structure  
A:Reference number: A30639; MUID:82189938  
A:Contents: annotation; x-ray crystallography, 2.0 angstroms  
R:Perwilliger, T.C.; Eisenberg, D.  
J. Biol. Chem. 257, 6016-6022, 1982  
A:Title: The structure of melittin. II. Interpretation of the structure.  
A:Reference number: A30640; MUID:82189959  
A:Contents: annotation; x-ray crystallography, 2.0 angstroms  
C:Superfamily: melittin major  
C:Keywords: amidated carboxyl end; blocked amino end; hemolysis; homotetramer; venom  
F:1-21/Domain: signal sequence #status experimental <SIG>  
F:22-43/Domain: propeptide #status experimental <PRO>  
F:44-69/Product: melittin major #status experimental <MAJ>  
F:51-69/Product: melittin F #status experimental <MEUF>  
F:44/Modified site: formulated amino end (Gly) (in mature form) (partial) #status experimental  
F:69/Modified site: amidated carboxyl end (Gln) (amide in mature form from following gly

Query Match 94.4%; Score 102; DB 1; Length 70;  
Best Local Similarity 100.0%; Pred. No. 4e-08;  
Matches 21; Conservative 0; Indels 0; Gaps 0;

QY 1 MKFLVNVVLFVWVVISYIYA 21  
|||||  
DB 1 MKFLVNVVLFVWVVISYIYA 21

```

I40841
immunodominant protein - heartwater rickettsia
C:Species: Cowdria ruminantium (heartwater rickettsia)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 29-Sep-1999
C:Accession: I40841
R:Mahan, S.M.; McGuire, T.C.; Semu, S.M.; Bowie, M.V.; Jongejan, F.; Rurangirwa, F.R.; F
Microbiology 140, 2135-2142, 1994
A:Title: Molecular cloning of a gene encoding the immunogenic 21kDa protein of Cowdria
A:Reference number: I40841; MUID:95005467
A:Accession: I40841
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-209 <RES>
A:Cross-references: GB:I07385; NID:q289922; PID:AAA50280.1; PID:q289923
C:Superfamily: immunodominant protein

Query Match 48.1% Score 52; DB 2; Length 209;
Best Local Similarity 35.0% Pred. No. 1.7;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKFLVNVALPFMWVYISYIY 20
:||||:|:| :|:| :|
DB 8 INFILNLCLEFAAIFLGYSY 27

RESULT 3
H69453
hypothetical protein AF1633 - Archaeoglobus fulgidus

```



C:Superfamily: Borrelia burgdorferi hypothetical protein BB0044

Query Match 45.4%; Score 49; DB 2; Length 133;  
Best Local Similarity 47.4%; Pred. No. 3.1;  
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 3 FLNVNVALFVMVYISYIYA 21  
||:| | :| | | :| | |

Db 12 FLISVFLIFVSGITFYFS 30

RESULT 6

NKVLHH core antigen - heron hepatitis virus

C:Species: heron hepatitis virus, HHV

A:Note: host Ardea cinerea (gray heron)

C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 16-Jul-1999

C:Accession: C30082

R:Sprengel, R.; Kaleta, E.F.; Will, H.

J. Virol. 62, 3832-3836, 1988

A:Title: Isolation and characterization of a hepatitis B virus endemic in herons.

A:Reference number: A93037; MUID:88333160

A:Accession: C30082

A:Molecule type: DNA

A:Residues: 1-305 <SPR>

A:Cross-references: GB:M22056; NID:G325452; PIDN:AAA45737.1; PID:G325453

C:Superfamily: hepatitis B virus core antigen

C:Keywords: core protein

Query Match 45.4%; Score 49; DB 1; Length 305;  
Best Local Similarity 55.0%; Pred. No. 6.1;  
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 FLNVNVALFVMVYISYIYAD 22

||| | | :| | | |

Db 26 FLVTPLVCTIVYDSCLYMD 45

RESULT 7

S44136

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - yeast (Candida parapsilosis) mito

C:Species: mitochondrion Candida parapsilosis

C:Date: 06-Feb-1995 #sequence\_revision 26-Jul-1996 #text\_change 07-Dec-1999

C:Accession: S44136

R:Nosek, J.

submitted to the EMBL Data Library, November 1993

A:Reference number: S44136

A:Accession: S44136

A:Molecule type: DNA

A:Residues: 1-40 <NOS>

A:Cross-references: EMBL:X75676; NID:G473032; PIDN:CAA53373.1; PID:G996061

C:Genetics:

A:Gene: ND2

A:Genome: mitochondrion

A:Genetic code: SGC3

C:Keywords: mitochondrion; NAD; oxidoreductase

Query Match 44.4%; Score 48; DB 2; Length 40;  
Best Local Similarity 35.0%; Pred. No. 1.6;  
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 MKFLNVNVALFVMVYISYIY 20

: | : | : | : | : |

Db 10 LAFVISLLLVFFITYFVIF 29

RESULT 8

S66287

sapacin B precursor - flesh fly (Sarcophaga peregrina)

C:Species: Sarcophaga peregrina

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Jul-1999  
C:Accession: S66287; S32323; JU0224

R:Lee, S.R.; Kurata, S.; Natori, S.

FEBS Lett. 368, 485-487, 1995

A:Title: Molecular cloning of cDNA for sapacin B, an antibacterial protein of Sarcoph

A:Reference number: S66287; MUID:95361929

A:Accession: S66287

A:Molecule type: mRNA

A:Residues: 1-88 <LEE>

A:Cross-references: GB:S80571; NID:G1245698; PIDN:AAB35004.1; PID:G1245699

R:Yamada, K.; Natori, S.

Biochem. J. 291, 275-279, 1993

A:Title: Purification, sequence and antibacterial activity of two novel sapacin homol

A:Reference number: S32323; MUID:93228618

A:Accession: S32323

A:Molecule type: protein

A:Residues: 55-88 <VAM>

C:Superfamily: sapacin precursor

C:Keywords: antibacterial; disulfide bond

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-54/Domain: propeptide #status predicted <PRO>

F:55-88/Product: sapacin B #status experimental <MAT>

F:57-78,64-84,68-86/Disulfide bonds: #status experimental

Query Match 44.4%; Score 48; DB 2; Length 88;

Best Local Similarity 47.4%; Pred. No. 3;

Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKFLNVNVALFVMVYISYI 19

||| : | : | : | :

Db 1 MKFLTSLLLLFVVMWSAV 19

RESULT 9

S62703

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 3 - Platymonas subcordiformis mito

C:Species: mitochondrion Platymonas subcordiformis

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 11-Jun-1999

C:Accession: S62703

R:Kessler, U.; Zetsche, K.

Plant Mol. Biol. 29, 1081-1086, 1995

A:Title: Physical map and gene organization of the mitochondrial genome from the unic

A:Reference number: S62702; MUID:96145517

A:Accession: S62703

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-117 <RES>

A:Cross-references: EMBL:Z47795; NID:G633584; PIDN:CAA87750.1; PID:G633586

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995

C:Genetics:

A:Gene: nad3

A:Genome: mitochondrion

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 3

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylati

Query Match 44.4%; Score 48; DB 2; Length 117;  
Best Local Similarity 36.8%; Pred. No. 3.9;  
Matches 7; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 3 FLNVNVALFVMVYISYIYA 21

||:| | :| | | :| | |

Db 11 FLFSLALASLIIFLSIFRA 29

RESULT 10

VXECSE

preprotein translocase secE chain - Escherichia coli

C:Species: Escherichia coli

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Jul-1999

C:Accession: A35139; A32873; H65204

R:Downing, W.L.; Sullivan, S.L.; Gottesman, M.E.; Dennis, P.P.

Mon Feb 4 15:23:47 2002

RESULT 12  
G69513  
conserved hypothetical protein AF2111 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1999  
C:Accession: G69513  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.C.; Gill, S.; Kirkness, E.  
Glodek, A.; Zhou, L.; Overbeek, R.; Cocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch  
A:Reference number: A69250; MUID:98049343  
A:Accession: G69513  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-213 <KLE>  
A:Cross-references: GB:AE000958; GB:AE000782; NID:g2689281; PIDN:AB89138.1; PID:g264  
C:Superfamily: conserved hypothetical protein MJ1677

Query Match 43.5%; Score 47; DB 2; Length 213;  
Best Local Similarity 47.8%; Pred. NO. 8.8;  
Matches 11; Conservative 6; Mismatches 4; Indels 2; Gaps 1;  
QY 1 MKFLV--NVALVFMVVIYSIYIA 21  
||| : ||| : ||| : ||| :  
DB 147 MKLLVLSAALVYSIVRLSHIYS 169

RESULT 13  
T26259  
hypothetical protein W07A8.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
C:Accession: T26259  
R:Basham, V.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z20184  
A:Accession: T26259  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-320 <WIL>  
A:Cross-references: EMBL:Z82075; PIDN:CAB04930.1; GSPDB:GN00023; CESP:W07A8.1  
A:Experimental source: clone W07A8  
C:Genetics:  
A:Gene: CESP:W07A8.1  
A:Map position: 5  
A:Introns: 83/1; 141/3; 256/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein W07A8.4

Query Match 41.7%; Score 45; DB 2; Length 320;  
Best Local Similarity 41.2%; Pred. NO. 24;  
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
QY 6 NVALVFMVVIYSIYIAD 22  
||| : ||| : ||| : ||| :  
DB 91 NVTLLVLCYLTPLFLD 107

RESULT 14  
T28208  
hypothetical protein ORF47 - Melanoplus sanguinipes entomopoxvirus  
C:Species: Melanoplus sanguinipes entomopoxvirus  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
C:Accession: T28208  
R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.  
J. Virol. 73, 533-552, 1999  
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.  
A:Reference number: Z20484; MUID:99102612  
A:Accession: T28208

Query Match 43.5%; Score 47; DB 1; Length 127;  
Best Local Similarity 40.9%; Pred. NO. 5.7;  
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
QY 1 MKFLVNVALLVFMVVIYSIYIAD 22  
||| : ||| : ||| : ||| :  
DB 17 MKWVVVALLVAVGNLYRD 38

RESULT 11  
E86089  
preprotein translocase [imported] - Escherichia coli (strain O157:H7)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
C:Accession: E86089  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: E86089  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-127 <STO>  
A:Cross-references: GB:AE005174; NID:g12518903; PIDN:AG59177.1; GSPDB:GN00145; UWGP:Z53  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: secE  
C:Superfamily: protein-export protein secE

Query Match 43.5%; Score 47; DB 2; Length 127;  
Best Local Similarity 40.9%; Pred. NO. 5.7;  
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
QY 1 MKFLVNVALLVFMVVIYSIYIAD 22  
||| : ||| : ||| : ||| :  
DB 17 MKWVVVALLVAVGNLYRD 38

Query Match 43.5%; Score 47; DB 2; Length 127;  
Best Local Similarity 40.9%; Pred. NO. 5.7;  
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
QY 1 MKFLVNVALLVFMVVIYSIYIAD 22  
||| : ||| : ||| : ||| :  
DB 17 MKWVVVALLVAVGNLYRD 38

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-331 <AFO>  
A:Cross-references: EMBL:AF063866; NID:g4049647; PIDN:AAC97829.1; PID:g4049869  
A:Experimental source: isolate Tuscon  
C:Genetics:  
A:Note: MSV047

Query Match 41.7%; Score 45; DB 2; Length 331;  
Best Local Similarity 31.6%; Pred. No. 25;  
Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 4 LVNVALVEMVYISYIYAD 22  
::: |::: | | | |  
Db 1 MINIKLIYLSLLISCVYTD 19

RESULT 15  
T15959  
hypothetical protein F02E8.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Feb-2000  
C:Accession: T15959  
R:Miller, N.  
submitted to the EMBL Data Library, April 1996  
A:Description: The sequence of C. elegans cosmid F02E8.  
A:Reference number: Z18436  
A:Accession: T15959  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-425 <MIL>  
A:Cross-references: EMBL:U53340; NID:g1255859; PID:g1255863; PIDN:AAA96209.1; GSPDB:GN00  
A:Experimental source: strain Bristol N2; clone F02E8  
C:Genetics:  
A:Gene: CESP:F02E8.2  
A:Map position: X  
A:Introns: 29/3; 69/2; 102/3; 153/1; 214/3; 273/1; 338/1; 379/1

Query Match 41.7%; Score 45; DB 2; Length 425;  
Best Local Similarity 50.0%; Pred. No. 31;  
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MKFLNVALVEMVYISY 18  
: | | | | | | | | | |  
Db 334 IKMLVTVTAVFFVCYLPY 351

Search completed: January 29, 2002, 11:00:41  
Job time: 2031 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:16:20 ; Search time 20.36 Seconds  
(without alignments)  
39.618 Million cell updates/sec

Title: US-09-763-397a-25

Perfect score: 108

Sequence: 1 MKFLVNVVFMVYIYAD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	102	94.4	70	1	MEL-APIME	P01501	apis mellif
2	51	47.2	227	1	YQ33-ARCFU	O28640	archaeoglob
3	49	45.4	133	1	Y044-BORBU	O51073	borrelia bu
4	49	45.4	305	1	CORA-HPBHE	P13845	heron hepat
5	48	44.4	88	1	SAPB-SARPE	P31529	sarcophaga
6	48	44.4	117	1	NU3M-PLASU	Q36518	platynomas
7	47	43.5	127	1	SRCE-ECOLI	P18920	escherichia
8	47	43.5	213	1	YL11-ARCFU	O28169	archaeoglob
9	46	42.6	677	1	SP87-DICDI	P54643	dictyosteli
10	45	41.7	1078	1	CASR-HUMAN	P41180	homo sapien
11	45	41.7	1079	1	CASR-MOUSE	O95966	mus musculu
12	45	41.7	1079	1	CASR-RAT	P48442	rattus norv
13	45	41.7	1085	1	CASR-BOVIN	P33384	bos taurus
14	44	40.7	218	1	YDPC-BACSU	P50738	bacillus su
15	44	40.7	271	1	Y223-AQUAE	O66414	aquifex aeo
16	44	40.7	295	1	SAPC-HAEIN	P45287	haemophilus
17	43	39.8	953	1	YMHG-YEAST	Q03516	saccharomyc
18	42.5	39.4	717	1	NAH4-RAT	O28739	rattus norv
19	42	38.9	278	1	YF33-ARCFU	O28739	archaeoglob
20	42	38.9	497	1	TH11-HUMAN	O60779	homo sapien
21	41.5	38.4	349	1	MRAY-CHLPN	Q92706	chlamydia p
22	40.5	37.5	533	1	MLO-HORVU	P93766	hordeum vul
23	40	37.0	128	1	NUOA-MYCTU	P95181	mycobacteri
24	40	37.0	181	1	LSPA-MYCGE	Q49401	mycoplasma
25	40	37.0	184	1	Y027-METJA	Q60336	methanococc
26	40	37.0	251	1	Y345-METJA	Q57965	methanococc
27	40	37.0	294	1	R1P1-TRIAN	P56626	trichosanth
28	40	37.0	397	1	CAPB-BACAN	P19580	bacillus an
29	40	37.0	401	1	NPT4-HUMAN	O00476	homo sapien
30	40	37.0	530	1	INR2-BOVIN	Q95141	bos taurus
31	40	37.0	586	1	CVDD-HAEIN	P45082	haemophilus
32	39.5	36.6	113	1	GUAU-PIG	O13009	sus scrofa
33	39.5	36.6	346	1	VGLM-HSV7J	P52372	human herpe

RESULT 1

ID	MEL-APIME	STANDARD;	PRT;	70 AA.
AC	P01501: P01503;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DE	MELITTIN PRECURSOR (ALLERGEN API M 3) (API M III).			
OS	Apis mellifera (Honeybee), and Apis cerana (Indian honeybee).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;			
OC	Apodea; Apidae; Apis;			
OX	NCBI_TaxID=7460, 7461;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=A.mellifera;			
RX	MEDLINE=83287387; PubMed=6309516;			
RA	Vlasak R., Unger-Ullmann C., Kreil G., Frischauf A.-M.;			
RT	"Nucleotide sequence of cloned cDNA coding for honeybee prepro-melittin.";			
RL	Eur. J. Biochem. 135:123-126(1983).			
RN	[2]			
RP	SEQUENCE OF 44-69 (MELITTIN 1 AND 2).			
RC	SPECIES=A.mellifera;			
RX	MEDLINE=68327913; PubMed=5592400;			
RA	Habermann E., Jentsch J.;			
RT	"Sequence analysis of melittin from tryptic and peptic degradation products.";			
RL	Hoppe-Seyler's Z. Physiol. Chem. 348:37-50(1967).			
RN	[3]			
RP	SYNTHESIS.			
RC	SPECIES=A.mellifera;			
RX	MEDLINE=72098668; PubMed=5139482;			
RA	Schroeder E., Luebke K., Lehmann M., Beetz I.;			
RT	"Haemolytic activity and action on the surface tension of aqueous solutions of synthetic melittins and their derivatives.";			
RL	Experientia 27:764-765(1971).			
RN	[4]			
RP	SYNTHESIS.			
RC	SPECIES=A.mellifera;			
RX	MEDLINE=72098669; PubMed=5139483;			
RA	Luebke K., Matthes S., Kloss G.;			
RT	"Isolation and structure of N1-formyl melittin.";			
RL	Experientia 27:765-767(1971).			
RN	[5]			
RP	SEQUENCE OF 44-69.			
RC	SPECIES=A.cerana;			
RA	Krell G.;			
RT	"Structure of melittin isolated from two species of honey bees.";			
RL	FEBS Lett. 33:241-244(1973).			
RN	[6]			
RP	X-RAY CRYSTALLOGRAPHY.			
RC	SPECIES=A.mellifera;			
RX	MEDLINE=82189959; PubMed=7076662;			
RA	Terwilliger T.C., Eisenberg D.;			
RT	"The structure of melittin. II. Interpretation of the structure.";			

096799 skeletonema  
016136 stomoxys ca  
P12983 vibrio algi  
P47759 mus musculu  
Q37626 prototheca  
P25871 nicotiana t  
P41852 aeromonas h  
Q37371 acanthamoeb  
Q95840 magnaporthe  
P03154 duck hepati  
Q09206 caenorhabdi  
Q9np09 homo sapien

```

RL J. Biol. Chem. 257:6016-6022(1982).
RN [7]
RP STRUCTURE BY NMR OF 44-69.
RC SPECIES=A.mellifera;
RA Barnham K.J., Hewish D., Werkmeister J., Curtin C., Kirkpatrick A.,
RA Bartone N., Norton R., Rivett D.;
RL Submitted (JUN-1998) to the PDB data bank.
RN [8]
RP REVIEW.
RX MEDLINE=90254148; PubMed=2187536;
RA Dempsey C.E.;
RT "The actions of melittin on membranes.";
RL Biochim. Biophys. Acta 1031:143-161(1990).
CC [-] FUNCTION: MAIN TOXIN OF BEE VENOM WITH STRONG HEMOLYTIC ACTIVITY.
CC INTEGRATES INTO CELL MEMBRANES AND HAS MULTIPLE EFFECTS, PROBABLY,
CC AS A RESULT OF ITS INTERACTION WITH NEGATIVELY CHARGED
CC PHOSPHOLIPIDS. IT INHIBITS WELL KNOWN TRANSPORT PUMPS SUCH AS THE
CC NA(+)-K(+)-ATPASE AND THE H(+)-K(+)-ATPASE. INCREASES THE
CC PERMEABILITY OF CELL MEMBRANES TO IONS, PARTICULARLY NA+ AND
CC INDIRECTLY CA2+, BECAUSE OF THE NA(+)-CA(2+)-EXCHANGE.
CC [-] SUBUNIT: MONOMER AND HOMOTETRAMER.
CC [-] MISCELLANEOUS: N-FORMYL-MELITTIN MAJOR HAS 80% OF THE ACTIVITY OF
CC MELITTIN.
CC [-] DATABASE: NAME=Protein Spotlight;
CC NOTE=Issue 12 of July 2001;
CC WWW="http://www.expasy.org/spotlight/articles/splitt012.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02007; CAA26038.1; -
DR PIR; A01761; MPBH1.
DR PIR; A01764; MEHB1.
DR PIR; A01762; MEHBC1.
DR PDB; 2MLT; 15-JUL-92.
DR PDB; 1BH1; 16-FEB-99.
DR InterPro; IPR002116; Melittin.
DR Pfam; PF01372; Melittin; 1.
DR Prodom; PDOM14636; Melittin; 1.
KW Hemolysis; Toxin; Venom; Amidation; Formylation; Signal; 3D-structure;
KW Allergen.
FT SIGNAL 1 21
FT PROPEP 22 43 REMOVED BY A DIPEPTIDYLPEPTIDASE.
FT CHAIN 44 69 MELITTIN.
FT MOD_RES 44 44 FORMYLATION (ABOUT 10% OF THE MOLECULES).
FT MOD_RES 69 69 AMIDATION (G-70 PROVIDE AMIDE GROUP).
FT VARIANT 64 64 K -> S (IN MELITTIN 2; POSSIBLY AN
FT ARTEFACT).
FT VARIANT 67 70 RQQG -> KRQQ (IN MELITTIN 2; POSSIBLY AN
FT ARTEFACT).
FT HELIX 45 53
FT TURN 54 54
FT HELIX 55 68
SQ SEQUENCE 70 AA; 7585 MW; 607F52C091C23BB6 CRC64;

Query Match 94.4%; Score 102; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 7.6e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLVNLVAFVWVVISYIVA 21
DB 1 MKFLVNLVAFVWVVISYIVA 21
|||||
RESULT 2
YG33_ARCFU STANDARD; PRT; 227 AA.
ID YG33_ARCFU
```

```

AC 028640;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE HYPOTHETICAL PROTEIN AF1633.
GN AF1633.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klek H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.D., Kervase A.R., Graham D.E., Kyte N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997)
CC [-] SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF000989; AAB89618.1; -
DR TIGR; AF1633; -
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 109 128 POTENTIAL.
FT TRANSMEM 173 192 POTENTIAL.
FT TRANSMEM 199 221 POTENTIAL.
SQ SEQUENCE 227 AA; 25313 MW; 63DCC7027C15A706 CRC64;

Query Match 47.2%; Score 51; DB 1; Length 227;
Best Local Similarity 42.9%; Pred. No. 2;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 MKFLVNLVAFVWVVISYIVA 21
DB 106 MKIMCVGLVLMVFAAFYFA 126
|||||
RESULT 3
Y044_BORBU STANDARD; PRT; 133 AA.
ID Y044_BORBU
AC 051073;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE HYPOTHETICAL PROTEIN BB0044.
GN BB0044.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
```

RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,  
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,  
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,  
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
RA Smith H.O., Venter J.C.;  
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
RT burgdorferi".  
RL Nature 390:580-586(1997).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE001118; AAC66442.1; -;  
DR TIGR; BB0044; -;  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 11 31 POTENTIAL.  
SQ SEQUENCE 133 AA; 16052 MW; A57686EA30F1959B CRC64;  
  
Query Match 45.4%; Score 49; DB 1; Length 133;  
Best Local Similarity 47.4%; Pred. No. 2.5;  
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
  
QY 3 FLVNVALVFMVYVYIYA 21  
DB 12 FLISVFLFIVSSGYFYS 30  
||:| |:| |:| |:|  
  
RESULT 4  
COR\_A\_HPBHE STANDARD; PRT; 305 AA.  
AC P13845;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 01-APR-1990 (Rel. 14, Last annotation update)  
DE CORE ANTIGEN.  
GN C.  
OS Heron hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.  
OX NCBI\_TaxID=28300;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88333160; PubMed=3418788;  
RA Sprengel R., Kaleta E.F., Will H.;  
RT "Isolation and characterization of a hepatitis B virus endemic in  
RT herons".  
RL J. Virol. 62:3832-3839(1988).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M22056; AAA45737.1; -;  
DR PIR; C30082; NKVLHH.  
DR InterPro; IPR002006; Hepatitis\_core.  
DR Pfam; PF00906; Hepatitis\_core; 1.  
KW Core protein.  
SQ SEQUENCE 305 AA; 34925 MW; 2B228BE44162CB06 CRC64;  
  
Query Match 45.4%; Score 49; DB 1; Length 305;  
Best Local Similarity 55.0%; Pred. No. 4.9;

Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
  
QY 3 FLVNVALVFMVYVYIYAD 22  
DB 26 FLVTVPVLTIVYDCLYMD 45  
||| |:| |:| |:| |:|  
  
RESULT 5  
SAPB\_SARPE STANDARD; PRT; 88 AA.  
AC P31529;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE SAPECIN B PRECURSOR.  
OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Sarcophagidae; Sarcophaga.  
OX NCBI\_TaxID=7386;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95361929; PubMed=7635204;  
RA Lee S.-R., Kurata S., Natori S.;  
RT "Molecular cloning of cDNA for sapecin B, an antibacterial protein of  
RT Sarcophaga, and its detection in larval brain".  
RL FEBS Lett. 368:485-487(1995).  
RN [2]  
RP SEQUENCE OF 55-88, AND DISULFIDE BONDS.  
RX MEDLINE=93228618; PubMed=8471044;  
RA Yamada K., Natori S.;  
RT "Purification, sequence and antibacterial activity of two novel  
RT sapecin homologues from Sarcophaga embryonic cells: similarity of  
RT sapecin B to charybdotoxin".  
RL Biochem. J. 291:275-279(1993).  
CC -!- FUNCTION: SAPECINS, WHICH ARE POTENT BACTERICIDAL PROTEINS,  
CC ARE PRODUCED IN RESPONSE TO INJURY. SAPECIN B IS CYTOTOXIC TO  
CC GRAM-POSITIVE BACTERIA.  
CC -!- TISSUE SPECIFICITY: HEMOCYTES AND FAT BODY.  
CC -!- INDUCTION: BY INJURY TO THE LARVAL CELL WALL.  
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; S80571; AAB35004.1; -;  
DR PIR; S32323; S32323.  
DR PIR; J00224; J00224.  
DR InterPro; IPR001542; Arthro\_defensin.  
DR Pfam; PF01097; Arthro\_defensin; 1.  
DR PROSITE; PS00425; ARTHROPOD\_DEFENSINS; FALSE\_NEG.  
KW Insect immunity; Antibiotic; Cleavage on pair of basic residues;  
KW Signal.  
FT SIGNAL 1 24 POTENTIAL.  
FT PROPEP 25 54  
FT PEPTIDE 55 88 SAPECIN B.  
FT DISULFID 57 78  
FT DISULFID 64 84  
FT DISULFID 68 86  
SQ SEQUENCE 88 AA; 10041 MW; 9459A0AF3B0E3D CRC64;  
  
Query Match 44.4%; Score 48; DB 1; Length 88;  
Best Local Similarity 47.4%; Pred. No. 2.5;  
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 MKFLVNVALVFMVYVYIYI 19  
|||| |:| |:| |:| |:|





```

CC CC -!- SIMILARITY: BELONGS TO THE SECE/SEC61-GAMMA FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; M30610; AAA24621.1; -
CC DR EMBL; U00006; AAC43079.1; -
CC DR EMBL; AAC00472; AAC76955.1; -
CC DR EMBL; AE005629; AAG59177.1; -
CC DR EMBL; AP002567; BAB38327.1; -
CC DR PIR; A35139; VXCSE.
CC DR PIR; SL6341; SL6341.
CC DR EcoGene; EG10939; secE.
CC DR InterPro; IPR001901; SecE.
CC DR Pfam; PF00584; SecE; 1.
CC DR PROSITE; PS01067; SECE_SEC61G; 1.
CC KW protein transport; Translocation; Transmembrane; Inner membrane;
CC KW Complete proteome.
CC FT DOMAIN 1 18 CYTOPLASMIC (PROBABLE).
CC FT TRANSMEM 19 36 PROBABLE.
CC FT DOMAIN 37 44 PERIPLASMIC (PROBABLE).
CC FT TRANSMEM 45 63 PROBABLE.
CC FT DOMAIN 64 92 CYTOPLASMIC (PROBABLE).
CC FT TRANSMEM 93 111 PROBABLE.
CC FT DOMAIN 112 127 PERIPLASMIC (PROBABLE).
CC SQ SEQUENCE 127 AA; 13643 MW; 94D37280522875CE CRC64;

```

```

Query Match 43.5%; Score 47; DB 1; Length 127;
Best Local Similarity 40.9%; Pred. No. 4.6;
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 MKFLVNVALLVGVNGLYRD 38
||:|||||:|:|:|:|
DB 17 MKWVVVALLVGVNGLYRD 38

RESULT 8
YL11_ARCFU
ID YL11_ARCFU STANDARD; PRT; 213 AA.
AC O28169;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN AF2111.
GN AF2111.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
CC Archaeoglobus.
CC NCBI_TaxID=2234;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton P.M., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).

```

```

CC CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC CC -!- SIMILARITY: BELONGS TO THE UPF0056 (MARC) FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; AE000958; AAB89138.1; -
CC DR TIGR; AF2111; -
CC DR InterPro; IPR002771; UPF0056.
CC DR Pfam; PF01914; UPF0056; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 1 21 POTENTIAL.
CC FT TRANSMEM 51 71 POTENTIAL.
CC FT TRANSMEM 75 95 POTENTIAL.
CC FT TRANSMEM 118 138 POTENTIAL.
CC FT TRANSMEM 142 162 POTENTIAL.
CC FT TRANSMEM 181 201 POTENTIAL.
CC SQ SEQUENCE 213 AA; 23062 MW; 5B1933795BC7278 CRC64;

Query Match 43.5%; Score 47; DB 1; Length 213;
Best Local Similarity 47.8%; Pred. No. 6.9;
Matches 11; Conservative 6; Mismatches 4; Indels 2; Gaps 1;

QY 1 MKFLV--NVALFMVVIYIYA 21
|||:|:|:|:|:|:|
DB 147 MKLLVLSAALVSVIRLSHIYS 169

RESULT 9
SP87_DICDI
ID SP87_DICDI STANDARD; PRT; 677 AA.
AC P54643;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SPORE COAT PROTEIN SP87 PRECURSOR (PL3 PROTEIN).
GN PSPD.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
CC NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=94229358; PubMed=8174787;
RA Yoder B.K., Mao J., Erdos G.W., West C.M., Blumberg D.D.;
RT "Identification of a new spore coat protein gene in the cellular
RT slime mold Dictyostelium discoideum."
RL Dev. Biol. 163:49-65(1994).
CC CC -!- FUNCTION: MAY CONTRIBUTE TO THE STRUCTURE OF THE COAT AT THE
CC CC INTERFACE BETWEEN THE MIDDLE, CELLULOSIC LAYER AND THE OUTER,
CC CC ELECTRON-DENSE, PROTEINOUS LAYER.
CC CC -!- SUBCELLULAR LOCATION: ACCUMULATES SPECIFICALLY IN REGULATED
CC CC SECRETORY VESICLES OF PRESPORE CELLS (PRESPORE VESICLES). THE
CC CC PROTEIN LATER ACCUMULATES EXTRACELLULARLY IN THE SPORE COAT.
CC CC -!- INDUCTION: BY C-AMP.
CC CC -!- PTM: DISULFIDE BONDING IS IMPORTANT FOR ASSOCIATING SP87 WITH THE
CC CC COAT.
CC CC -!- PTM: PHOSPHORYLATED ON SERINE RESIDUES.
CC CC -!- SIMILARITY: CONTAINS 12 PRESPORE MOTIFS.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: U25144; AAA73515.1;
DR HSP: P06620; IINA.
DR Dictydb: DD02054; psp.
DR InterPro: IPR003645; FOLN.
DR SMART: SM00274; FOLN: 5.
DR Glycoprotein; Repeat; Sporulation; Signal.
FT STIGNAL 1 21
FT CHAIN 22 677
FT REPEAT 189 201 SPORE COAT PROTEIN SP87.
FT REPEAT 232 244 PRESPORE MOTIF 1.
FT REPEAT 232 244 PRESPORE MOTIF 2.
FT REPEAT 276 288 PRESPORE MOTIF 3.
FT REPEAT 316 328 PRESPORE MOTIF 4.
FT REPEAT 356 368 PRESPORE MOTIF 5.
FT REPEAT 390 402 PRESPORE MOTIF 6.
FT REPEAT 418 430 PRESPORE MOTIF 7.
FT REPEAT 450 462 PRESPORE MOTIF 8.
FT REPEAT 513 525 PRESPORE MOTIF 9.
FT REPEAT 543 555 PRESPORE MOTIF 10.
FT REPEAT 571 583 PRESPORE MOTIF 11.
FT REPEAT 601 613 PRESPORE MOTIF 12.
FT DOMAIN 148 192 5 X 9 AA REPEATS OF G-G-S-S-S-G-G-T-S.
FT REPEAT 148 156 1-1 (APPROXIMATE).
FT REPEAT 157 165 1-2.
FT REPEAT 166 174 1-3.
FT REPEAT 175 183 1-4 (APPROXIMATE).
FT REPEAT 184 192 1-5 (APPROXIMATE).
FT DOMAIN 268 387 3 X 40 AA APPROXIMATE REPEATS.
FT REPEAT 268 307 2-1.
FT REPEAT 308 347 2-2.
FT REPEAT 348 387 2-3.
FT DOMAIN 391 628 8 X 26 AA APPROXIMATE REPEATS.
FT REPEAT 391 417 3-1.
FT REPEAT 419 445 3-2.
FT REPEAT 451 477 3-3.
FT REPEAT 480 506 3-4.
FT REPEAT 514 540 3-5.
FT REPEAT 544 570 3-6.
FT REPEAT 572 598 3-7.
FT REPEAT 602 628 3-8.
FT DOMAIN 141 185 GLY/SER/THR-RICH.
FT DOMAIN 256 351 GLY/SER/THR-RICH.
FT DOMAIN 342 371 GLY/SER/THR-RICH.
FT DOMAIN 174 183 POLY-SER.
FT DOMAIN 445 448 POLY-PRO.
FT CARBOHYD 72 72 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 677 AA; 71804 MW; EF6E29CF57E78D5 CRC64;

Query Match 42.68; Score 46; DB 1; Length 677;
Best Local Similarity 42.98; Pred. No. 24;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKELVNVLFVWVVISIYA 21.
DB 1 MFLKNGVFFWFLVSKSYA 21

RESULT 10
CASR_HUMAN STANDARD; PRT; 1078 AA.
AC P41180; Q13912; Q16379; Q16108; Q16110;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE EXTRACELLULAR CALCIUM-SENSING RECEPTOR PRECURSOR (CASR) (PARATHYROID
DE CELL CALCIUM-SENSING RECEPTOR).
GN CASR OR GPRC2A OR PCAR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Pearce S.H.S., Thakker R.V.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Parathyroid;
RX MEDLINE=95279439; PubMed=7759551;
RA Garrett J.E., Capuano I.V., Hammerland L.G., Hung B.C., Brown E.M.,
RA Hebert S.C., Nemeth E.F., Fuller F.;
RT "Molecular cloning and functional expression of human parathyroid
calcium receptor cDNAs.";
RL J. Biol. Chem. 270:12919-12925(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95408281; PubMed=7677761;
RA Aida K., Koishi S., Tawata M., Onaya T.;
RT "Molecular cloning of a putative Ca(2+)-sensing receptor cDNA from
human kidney.";
RL Biochem. Biophys. Res. Commun. 214:524-529(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=96343806; PubMed=8756555;
RA Frelch M., Zink-Lorenz A., Holloschi A., Hafner M., Flockerzi V.,
RA Raue F.;
RT "Expression of a calcium-sensing receptor in a human medullary
thyroid carcinoma cell line and its contribution to calcitonin
secretion.";
RL Endocrinology 137:3842-3848(1996).
RN [5]
RP SEQUENCE OF 643-908 FROM N.A.
RX MEDLINE=96193893; PubMed=8613532;
RA Bikle D.D., Ratnam A., Mauro T., Harris J., Pillai S.;
RT "Changes in calcium responsiveness and handling during keratinocyte
differentiation. Potential role of the calcium receptor.";
RL J. Clin. Invest. 97:1085-1093(1996).
RN [6]
RP VARIANTS FHH GLN-185; LYS-297 AND TRP-795.
RX MEDLINE=94094324; PubMed=7916660;
RA Pollak M.R., Brown E.M., Chou Y.-H.W., Hebert S.C., Marx S.J.,
RA Steinmann B., Levi T., Seidman C.E., Seidman J.G.;
RT "Mutations in the human Ca(2+)-sensing receptor gene cause familial
hypocalcemic hypercalcaemia and neonatal severe
hyperparathyroidism.";
RL Cell 75:1297-1303(1993).
RN [7]
RP VARIANT ADH ALA-127.
RX MEDLINE=95179179; PubMed=7874174;
RA Pollak M.R., Brown E.M., Estep H.L., McLaine P.N., Kifor O., Park J.,
RA Hebert S.C., Seidman C.E., Seidman J.G.;
RT "Autosomal dominant hypocalcaemia caused by a Ca(2+)-sensing receptor
gene mutation.";
RL Nat. Genet. 8:303-307(1994).
RN [8]
RP VARIANTS FHH MET-62; CYS-66; MET-138; GLU-143 AND GLN-227.
RX MEDLINE=95243222; PubMed=7726161;
RA Chou Y.-H.W., Pollak M.R., Brandl M.L., Toss G., Aronqvist H.,
RA Atkinson A.B., Papapoulos S.E., Marx S., Brown E.M., Seidman J.G.,
RA Seidman C.E.;
RT "Mutations in the human Ca(2+)-sensing-receptor gene that cause
familial hypocalcemic hypercalcaemia.";
RL Am. J. Hum. Genet. 56:1075-1079(1995).
RN [9]
RP SEQUENCE OF 1-61 FROM N.A., AND VARIANT FHH ALA-39.
RX MEDLINE=95403641; PubMed=7673400;
RA Aida K., Koishi S., Inoue M., Nakazato M., Tawata M., Onaya T.;
RT "Familial hypocalcemic hypercalcaemia associated with mutation in the
human Ca(2+)-sensing receptor gene.";
RL J. Clin. Endocrinol. Metab. 80:2594-2598(1995).
RN [10]
RP VARIANTS NSHPT LEU-227 AND TYR-598.
RX MEDLINE=96292293; PubMed=8675635;
RA Pearce S.H.S., Trump D., Woodling C., Besser G.M., Chew S.L.,
RA

```

RA Grant D.B., Heath D.A., Hughes I.A., Paterson C.R., Whyte M.P.,  
 RA Thakker R.V.;  
 RT "Calcium-sensing receptor mutations in familial benign hypercalcaemia  
 RT and neonatal hyperparathyroidism.";  
 RL J. Clin. Invest. 96:2683-2692(1995).  
 RN [11]  
 RP VARIANTS ADHP THR-116; HIS-681 AND SER-806, AND VARIANT SER-851.  
 RX MEDLINE=96311554; PubMed=8733126;  
 RA Baron J., Winer K.K., Yanovski J.A., Cunningham A.W., Laue L.,  
 RA Zimmerman D., Cutler G.B. Jr.;  
 RT "Mutations in the Ca(2+)-sensing receptor gene cause autosomal  
 RT dominant and sporadic hypoparathyroidism.";  
 RL Hum. Mol. Genet. 5:601-606(1996).  
 RN [12]  
 RP VARIANT FHH ARG-174.  
 RX MEDLINE=97442275; PubMed=9298824;  
 RA Ward B.K., Stuckey B.G.A., Gutteridge D.H., Laing N.G., Pullan P.T.,  
 RA Ratajczak T.;  
 RT "A novel mutation (L174R) in the Ca2+-sensing receptor gene  
 RT associated with familial hypocalcaemic hypercalcaemia.";  
 RL Hum. Mutat. 10:233-235(1997).  
 CC -!- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF  
 CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A  
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND  
 CC MESSENGER SYSTEM.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS SEEM TO BE PRODUCED BY  
 CC ALTERNATIVE SPLICING.  
 CC -!- TISSUE SPECIFICITY: FOUND IN KIDNEY, BUT NOT IN BRAIN, LUNG,  
 CC LIVER, HEART, SKELETAL MUSCLE, OR PLACENTA.  
 CC -!- DISEASE: DEFECTS IN CASR ARE A CAUSE OF FAMILIAL HYPOCALCAEMIC  
 CC HYPERCALCAEMIA (FHH) AND NEONATAL SEVERE HYPERPARATHYROIDISM  
 CC (NSHPT). TWO INHERITED CONDITIONS CHARACTERIZED BY ALTERED CALCIUM  
 CC HOMEOSTASIS. THE MUTATIONS REDUCE THE ACTIVITY OF THE RECEPTOR.  
 CC FHH AFFECTED INDIVIDUALS EXHIBIT MILD OR MODEST HYPERCALCAEMIA,  
 CC RELATIVE HYPOCALCAEMIA, AND INAPPROPRIATELY NORMAL PTH LEVELS. IN  
 CC CONTRAST, NSHPT IS A RARE AUTOSOMAL RECESSIVE LIFE-THREATENING  
 CC DISORDER CHARACTERIZED BY VERY HIGH SERUM CALCIUM CONCENTRATIONS,  
 CC SKELETAL DEMINERALIZATION, AND PARATHYROID HYPERPLASIA. IN SOME  
 CC INSTANCES NSHPT HAS BEEN DEMONSTRATED TO BE THE HOMOZYGOUS FORM OF  
 CC FHH.  
 CC -!- DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT  
 CC HYPOCALCAEMIA (ADHP) IN WHICH THE RECEPTOR IS ACTIVATED AT SUBNORMAL  
 CC CA(2+) LEVELS.  
 CC -!- DISEASE: DEFECTS IN PCAR1 ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT  
 CC HYPOPARATHYROIDISM (ADHP). ADHP IS CHARACTERIZED BY HYPOCALCAEMIA  
 CC AND HYPERPHOSPHATEMIA DUE TO INADEQUATE SECRETION OF PARATHYROID  
 CC HORMONE. SYMPTOMS ARE SEIZURES, TETANY AND CRAMPS.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X81086; CAA56990.1; -  
 DR EMBL; U20759; AAA86503.1; -  
 DR EMBL; U20760; AAA86504.1; -  
 DR EMBL; D50855; BAA09453.1; -  
 DR EMBL; S83176; AAB46873.1; -  
 DR EMBL; S79217; AAB35262.2; -  
 DR EMBL; S68032; AAB29413.2; ALT\_SEQ.  
 DR EMBL; S68033; AAB29414.1; -  
 DR EMBL; S68036; AAB29415.1; -  
 DR EMBL; S81755; AAD14370.1; -  
 DR GCRDB; GCR\_1337; -  
 DR GCRDB; GCR\_1874; -  
 DR GCRDB; GCR\_2012; -  
 DR GCRDB; GCR\_2013; -  
 DR GCRDB; GCR\_2696; -

DR GCRDB; GCR\_2697; -  
 DR MIM; 601199; -  
 DR MIM; 145980; -  
 DR MIM; 601198; -  
 DR InterPro; IPR001828; ANF\_receptor.  
 DR Pfam; PF00003; 7tm\_3; 1.  
 DR PRINTS; PR00248; GPCR\_MGR.  
 DR PROSITE; PS00592; CASSENSING.  
 DR PROSITE; PS00979; G\_PROTEIN\_RECEP\_F3\_1; 1.  
 DR PROSITE; PS00980; G\_PROTEIN\_RECEP\_F3\_2; 1.  
 DR PROSITE; PS00981; G\_PROTEIN\_RECEP\_F3\_3; 1.  
 DR PROSITE; PS02559; G\_PROTEIN\_RECEP\_F3\_4; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 KW Disease mutation; Alternative splicing; Polymorphism.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 1078 EXTRACELLULAR CALCIUM-SENSING RECEPTOR.  
 FT DOMAIN 20 612 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 613 635 I (POTENTIAL).  
 FT DOMAIN 636 649 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 650 670 II (POTENTIAL).  
 FT DOMAIN 671 681 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 682 700 III (POTENTIAL).  
 FT DOMAIN 701 724 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 725 745 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 746 769 IV (POTENTIAL).  
 FT TRANSMEM 770 792 V (POTENTIAL).  
 FT DOMAIN 793 805 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 806 828 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 829 836 VI (POTENTIAL).  
 FT TRANSMEM 837 862 VII (POTENTIAL).  
 FT DOMAIN 863 1078 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 QY 2 KFLNVNVALFMVYISYIYA 21  
 DB 805 KETFSMLFFIWIISFIPA 824  
 RESULT 11  
 ID CASR\_MOUSE STANDARD; PRT; 1079 AA.  
 AC Q9QY96; O08968; O88519; Q9QY95; Q9QZU8; Q9R1D6; Q9R1Y2;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE EXTRACELLULAR CALCIUM-SENSING RECEPTOR PRECURSOR (CASR) (PARATHYROID  
 DE CELL CALCIUM-SENSING RECEPTOR).  
 GN CASR OR GPRC2A.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.  
 RC STRAIN=C57BL/6; TISSUE=Kidney;  
 RX MEDLINE=20092890; PubMed=10625662;  
 RA Oda Y., Tu C.-L., Chang W., Crumrine D., Koemueves L., Mauro T.,  
 RA Elias P.M., Birk D.D.;  
 RT "The calcium sensing receptor and its alternatively spliced form in  
 RT murine epidermal differentiation.";

Query Match 41.7%; Score 45; DB 1; Length 1078;  
 Best Local Similarity 45.0%; Pred. No. 47;  
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;





Db 37 FELGVVLVFPIMEIQV 53

RESULT 15

Y223\_AQUAE STANDARD; PRT; 271 AA.

ID Y223\_AQUAE

AC O66414;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE HYPOTHETICAL PROTEIN AA23.

GN AA23.

OS Aquifex aeolicus.

OG Bacterioidetes.

OC Bacteria; Aquificales; Aquificaceae; Aquifex.

OX NCBI\_TaxID=63363;

RP SEQUENCE FROM N.A.

RC STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Huber R.,

RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujaay M., Swanson R.V.;

RA Feidman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex

RL Nature 392:353-358(1998).

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; AF000667; AAC07966.1; -

KW Hypothetical protein; Transmembrane; Plasmid; Complete proteome.

FT TRANSMEM 30 50 POTENTIAL.

FT TRANSMEM 189 209 POTENTIAL.

FT TRANSMEM 218 238 POTENTIAL.

SQ SEQUENCE 271 AA; 31365 MW; 6CE96A93A61953DF CRC64;

Query Match 40.7%; Score 44; DB 1; Length 271;

Best Local Similarity 27.3%; Pred. No. 22;

Matches 6; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 MKFLVNVLFVFMVYISYIYAD 22

Db 226 LTYLSNIPMIFACWMDHFVAD 247

Search completed: January 29, 2002, 11:16:21

Job time: 150 sec

Query Match 40.7%; Score 44; DB 1; Length 218;

Best Local Similarity 41.2%; Pred. No. 18;

Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 FLVNVLFVFMVYISYI 19

Db 806 KFIIFSMLFFIVWISFIPA 825

RESULT 14

YPDC\_BACSU STANDARD; PRT; 218 AA.

ID YPDC\_BACSU

AC P50738;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE HYPOTHETICAL 24.7 KDA PROTEIN IN RECO-CMK INTERGENIC REGION.

GN YPDC.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI\_TaxID=1423;

RP SEQUENCE FROM N.A.

RX STRAIN=168 / MARBURG;

RX MEDLINE=96349105; PubMed=8760912;

RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,

RA Serron P.;

RT "Sequence analysis of the Bacillus subtilis chromosome region between

RT the serA and kds loci cloned in a yeast artificial chromosome.";

RL Microbiology 142:2005-2016(1996)

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; LA7648; AAC83956.1; -

DR EMBL; Z99115; CAB14210.1; -

DR EMBL; Z99116; CAB14226.1; -

DR Subtilist; BG11438; Ypdc.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 30 50 POTENTIAL.

FT TRANSMEM 65 85 POTENTIAL.

FT TRANSMEM 127 147 POTENTIAL.

FT TRANSMEM 180 200 POTENTIAL.

SQ SEQUENCE 218 AA; 24719 MW; 88556D50863E14BC CRC64;

Query Match 40.7%; Score 44; DB 1; Length 218;

Best Local Similarity 41.2%; Pred. No. 18;

Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 FLVNVLFVFMVYISYI 19

Db 806 KFIIFSMLFFIVWISFIPA 825

RESULT 15

Y223\_AQUAE STANDARD; PRT; 271 AA.

ID Y223\_AQUAE

AC O66414;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE HYPOTHETICAL PROTEIN AA23.

GN AA23.

OS Aquifex aeolicus.

OG Bacterioidetes.

OC Bacteria; Aquificales; Aquificaceae; Aquifex.

OX NCBI\_TaxID=63363;

RP SEQUENCE FROM N.A.

RC STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Huber R.,

RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujaay M., Swanson R.V.;

RA Feidman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex

RL Nature 392:353-358(1998).

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; AF000667; AAC07966.1; -

KW Hypothetical protein; Transmembrane; Plasmid; Complete proteome.

FT TRANSMEM 30 50 POTENTIAL.

FT TRANSMEM 189 209 POTENTIAL.

FT TRANSMEM 218 238 POTENTIAL.

SQ SEQUENCE 271 AA; 31365 MW; 6CE96A93A61953DF CRC64;

Query Match 40.7%; Score 44; DB 1; Length 271;

Best Local Similarity 27.3%; Pred. No. 22;

Matches 6; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 MKFLVNVLFVFMVYISYIYAD 22

Db 226 LTYLSNIPMIFACWMDHFVAD 247

Search completed: January 29, 2002, 11:16:21

Job time: 150 sec

Query Match 40.7%; Score 44; DB 1; Length 218;

Best Local Similarity 41.2%; Pred. No. 18;

Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 FLVNVLFVFMVYISYI 19

Db 806 KFIIFSMLFFIVWISFIPA 825



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:59:20 ; Search time 66.28 Seconds  
(without alignments)  
24.587 Million cell updates/sec

Title: US-09-763-397A-25

Perfect score: 108

Sequence: 1 MKFLVNLVFMVYISYIYAD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	108	100.0	22	21	AA1980
2	108	100.0	350	21	AA1981
3	108	100.0	354	21	AA1982
4	108	100.0	354	21	AA1983
5	108	100.0	354	21	AA1984
6	104	96.3	386	22	AA2001
7	102	94.4	21	22	AA1985
8	102	94.4	49	21	AA1986
9	102	94.4	177	21	AA1987
10	97	89.8	21	17	AA1988
11	97	89.8	21	20	AA1989

12	93.5	86.6	26	19	AAW64613	pMelBac secretion
13	93.5	86.6	26	19	AAW61590	Targeting molecule
14	93.5	86.6	26	19	AAW85766	Secretion signal a
15	55	50.9	205	19	AAW51098	Ehrlichia canis ma
16	55	50.9	205	21	AAW36192	Ehrlichia canis MA
17	55	50.9	205	22	AAU04202	Major antigenic pr
18	54	50.0	205	19	AAW51099	Ehrlichia chaffeen
19	54	50.0	205	21	AAW36193	Ehrlichia chaffeen
20	54	50.0	205	22	AAU04203	Major antigenic pr
21	52	48.1	205	21	AAW36194	Cowdria ruminatium
22	49	45.4	133	20	AAW19802	B. burgdorferi ant
23	49	45.4	305	20	AAW29763	Human hepatitis B
24	49	45.4	397	19	AAW80573	Human O-fucosyltra
25	47	43.5	206	18	AAW20456	H. pylori flagella
26	45	41.7	40	20	AAW02715	Human secreted pro
27	45	41.7	250	20	AAW49110	Human CaR transmem
28	45	41.7	425	22	AAU03853	G protein-coupled
29	45	41.7	901	21	AAW45001	Human calcium sens
30	45	41.7	974	21	AAW45000	Human calcium sens
31	45	41.7	1001	21	AAW44999	Human calcium sens
32	45	41.7	1026	18	AAW32059	Dogfish shark kidn
33	45	41.7	1058	18	AAW25762	Amino acid sequenc
34	45	41.7	1059	22	AAU00508	Chicken calcium-se
35	45	41.7	1078	17	AAW11889	Parathyroid calculi
36	45	41.7	1078	19	AAW54846	Human parathyroid
37	45	41.7	1078	19	AAW38274	Human parathyroid
38	45	41.7	1078	20	AAW28840	Human calcium rece
39	45	41.7	1078	20	AAW41780	Human parathyroid
40	45	41.7	1078	20	AAW89565	Human parathyroid
41	45	41.7	1078	21	AAW51827	Human calcium rece
42	45	41.7	1078	21	AAW70325	Human wild type ca
43	45	41.7	1078	22	AAU02195	Cynomolgous monkey
44	45	41.7	1078	22	AAW74391	Protein encoded by
45	45	41.7	1079	19	AAW54847	Rat kidney calcium

ALIGNMENTS

RESULT 1

AAW70301

ID AAW70301 standard; peptide: 22 AA.

XX AC AAW70301;

XX DT 06-JUN-2000 (first entry)

XX XX Honey bee melittin signal peptide.

XX DE Recombinant protein; CDC/NIH/VAC-1; multivalent; malaria; vaccine;  
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment; SSP-2;  
KW circumsporozoite protein; CSP; sporozoite surface protein-2; MSP-1; MSP-2;  
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;  
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;  
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;  
KW Pf27; antiparasitic; prevention; anti-CDC/NIH/VAC-1 antibody;  
KW melittin; honey bee.

XX OS Apis sp.

XX XX WO200011179-A1.

XX XX 02-MAR-2000.

XX XX 19-AUG-1999; 99WO-US18869.

XX XX 21-AUG-1998; 98US-0097703.

XX XX (NAIM-) NAT INST IMMUNOLOGY.  
XX XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Lal AA, Shi YP, Hasnain SE;

DR WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection

XX comprises antigenic peptides obtained from different stages of

PT plasmodium falciparum life cycle -

XX Claim 2; Page 16; 52pp; English.

XX The present sequence is the melittin secretory signal peptide from

CC honey bee. It is used in the construction of recombinant protein

CC CDC/NIIMALVAC-1, which is a multivalent, multistage malarial

CC vaccine. The recombinant protein comprises, melittin signal peptide,

CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes

CC from circumsporozoite protein (CSP), sporozoite surface protein-2

CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1

CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding

CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete

CC specific antigen, Pf27. These epitopes were obtained at different stages

CC of the life cycle of P. falciparum. CDC/NIIMALVAC-1 vaccine has

CC antiparasitic activity and can be used for treatment and prevention of

CC malarial infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for

CC detecting P. falciparum in biological samples.

XX Sequence 22 AA;

XX Query Match 100.0%; Score 108; DB 21; Length 22;

XX Best Local Similarity 100.0%; Pred. No. 1.6e-10;

XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLNVNVALFVWVYISYIYAD 22

DB 1 mkflnvvalvfmvvyisyiyad 22

RESULT 2

AA170278

ID AAY70278 standard; Protein; 350 AA.

XX AC AAY70278;

XX 06-JUN-2000 (first entry)

XX Recombinant vaccine CDC/NIIMALVAC-1.

XX Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;

KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;

KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;

KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;

KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;

KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;

KW Pf27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody;

KW honey bee.

XX Chimeric - Apis sp.

OS Chimeric - Clostridium tetani.

OS Chimeric - Plasmodium falciparum.

XX Key Location/Qualifiers

FH 1..22

FT /label= Melittin\_signal\_peptide

FT /note= "Derived from Honey bee"

FT 23..350

FT /label= Mature\_CDC/NIIMALVAC-1

FT /note= "Recombinant multivalent malarial vaccine"

XX WO200011179-A1.

XX 02-MAR-2000.

XX 19-AUG-1999; 99WO-US18869.

XX 21-AUG-1998; 98US-0097703.

XX (NAIM-) NAT INST IMMUNOLOGY.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Lal AA, Shi YP, Hasnain SE;

XX WPI; 2000-237654/20.

XX N-PSDB; AAZ51336.

XX Novel recombinant protein as vaccine for treating malarial infection

XX comprises antigenic peptides obtained from different stages of

PT plasmodium falciparum life cycle -

XX Claim 3; Page 43-44; 52pp; English.

XX The present sequence is that of recombinant protein CDC/NIIMALVAC-1,

CC which is a multivalent, multistage malarial vaccine. The recombinant

CC protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope

CC from tetanus toxoid and 21 antigenic epitopes from circumsporozoite

CC protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage

CC antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical

CC membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),

CC rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Pf27.

CC These epitopes were obtained at different stages of the life cycle of

CC Plasmodium falciparum. CDC/NIIMALVAC-1 vaccine has antiparasitic

CC activity and can be used for treatment and prevention of malarial

CC infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for detecting

CC P. falciparum in biological samples.

XX Sequence 350 AA;

XX Query Match 100.0%; Score 108; DB 21; Length 350;

XX Best Local Similarity 100.0%; Pred. No. 2.8e-09;

XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLNVNVALFVWVYISYIYAD 22

DB 1 mkflnvvalvfmvvyisyiyad 22

RESULT 3

AA10640

ID AAB10640 standard; Protein; 354 AA.

XX AC AAB10640;

XX 19-JAN-2001 (first entry)

XX Human VEGF-X protein for expression in Baculovirus/insect cell systems.

XX VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;

XX antineumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;

XX angiogenesis regulator; vascularization regulator; cancer; psoriasis;

XX rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;

XX tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;

XX venous sore; diabetic ulcer; burns; skin graft growth.

XX Homo sapiens.

XX WO200037641-A2.

XX 29-JUN-2000.

XX 21-DEC-1999; 99WO-US30503.

XX 22-DEC-1998; 98GB-0028377.

XX 18-MAR-1999; 99US-0124967.

XX 08-NOV-1999; 99US-0164131.

XX (JANC ) JANSSEN PHARM NV.

XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;

PI



PR 14-AUG-1997; 97US-0911312.  
 PR 14-AUG-1997; 97US-0912951.  
 XX (GERO-) GERON CORP.  
 PA (UYTE-) UNIV TECHNOLOGY CORP.  
 XX Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J;  
 PI Morin GB, Nakamura T, Harley CB;  
 XX WPI; 1998-171633/16.  
 XX Pure and recombinant human Telomerase Reverse Transcriptase and its  
 PT variants - are useful in the diagnosis, prognosis and treatment of  
 PT cell proliferation conditions especially cancer and ageing  
 XX Example 6; Page 234-235; 387pp; English.  
 XX The present sequence represents a fusion protein from an example  
 CC of the present invention which describes human telomerase reverse  
 CC transcriptase (hTERT). The present invention also describes the  
 CC following methods: (A) determining whether a test compound is  
 CC a modulator of hTERT, by detecting the change in hTERT recombinant  
 CC protein or polynucleotide, on administration of the compound;  
 CC (B) preparation of recombinant telomerase by contacting a protein  
 CC preparation of hTERT with a telomerase RNA component; (C) detection  
 CC of the hTERT RNA or protein in a sample by binding a relevant  
 CC probe to the sample and detecting the complex formed or in the case of  
 CC RNA detection, amplifying the product and correlating the presence of  
 CC complex or amplification product with presence of hTERT in the sample;  
 CC and (D) increasing the proliferation of a vertebrate cell by increasing  
 CC hTERT expression; and (E) the use of an agent that causes an increase in  
 CC cell vertebrate cell proliferation to create a medicament that inhibits  
 CC ageing. A protein preparation of hTERT and the polynucleotide encoding  
 CC hTERT can be used in the manufacture of medicaments for inhibiting the  
 CC effect of ageing or cancer. Inhibitors of telomerase activity can be  
 CC used to treat conditions that are associated with high telomerase  
 CC activity. A protein preparation of hTERT can also be used in the new  
 CC methods.  
 XX Sequence 1189 AA;  
 SQ

Query Match 100.0%; Score 108; DB 19; Length 1189;  
 Best Local Similarity 100.0%; Pred. No. 1e-08;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKELVNVALVFMVYISYIYAD 22  
 |||||  
 DB 1 mkflvnvalvmvvyisyiyad 22

RESULT 6  
 AAE05128  
 ID AAE05128 standard; Protein; 386 AA.  
 AC AAE05128;  
 XX 18-SEP-2001 (first entry)  
 DT Infectious salmon anaemia virus (ISAV) 92-M fusion protein.  
 DE Infectious salmon anaemia virus; ISAV; 92-M clone; vaccine; prophylaxis;  
 KW infectious salmon anaemia; therapy; fusion protein; antianaemic.  
 KW Chimeric - Infectious salmon anaemia virus.  
 OS Chimeric - Unidentified.  
 XX Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /label= Mellitin\_signal\_peptide  
 FT Peptide 22..26  
 FT /label= Linker\_peptide  
 FT Protein 27..386

FT /note= "Mature ISAV 92-M protein"  
 XX WO200149712-A2.  
 PN 12-JUL-2001.  
 PD 03-JAN-2001; 2001WO-EP00046.  
 XX 07-JAN-2000; 2000EP-0200054.  
 XX 29-FEB-2000; 2000EP-0200700.  
 PR (ALKU ) AKZO NOBEL NV.  
 PA Biering E, Krossoy B;  
 XX WPI; 2001-441845/47.  
 DR N-PSDB; AAD09875.  
 XX Novel vaccine for treatment and/or prevention of infectious salmon  
 PT anemia in fish -  
 PT Claim 7; Page 39-41; 43pp; English.  
 XX The present invention relates to vaccine for prevention and/or  
 CC prophylaxis of infectious salmon anaemia in fish. The invention  
 CC provides for nucleic acid sequence encoding viral proteins of  
 CC infectious salmon anaemia virus (ISAV) as well as the isolated protein.  
 CC Nucleic acids encoding viral protein is useful for the manufacture of a  
 CC DNA vaccine for diagnosis, treatment and/or prophylaxis of infectious  
 CC salmon anaemia in fish, and viral protein can be used for the  
 CC manufacture of antibodies that are specific for ISAV. The present  
 CC sequence is infectious salmon anaemia virus 92-M fusion protein.  
 XX Sequence 386 AA;  
 SQ

Query Match 96.3%; Score 104; DB 22; Length 386;  
 Best Local Similarity 95.5%; Pred. No. 1.3e-08;  
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKFLVNVALVFMVYISYIYAD 22  
 |||||  
 DB 1 mkflvnvalvmvvyisyiyad 22

RESULT 7  
 AAB72437  
 ID AAB72437 standard; Peptide; 21 AA.  
 XX AC AAB72437;  
 XX 08-MAY-2001 (first entry)  
 DT Melittin signal peptide.  
 DE UGGT; UDP-glucose:glycoprotein glucosyltransferase; enzyme; honeybee;  
 KW melittin signal peptide.  
 KW Apis mellifica.  
 OS WO200112845-A1.  
 PN 22-FEB-2001.  
 PD 27-JUL-2000; 2000WO-CA00883.  
 XX 18-AUG-1999; 99US-0376330.  
 XX (CANA ) NAT RES COUNCIL CANADA.  
 PA Tessier DC, Dignard D, Bergeron JJM, Thomas DY;  
 PI WPI; 2001-218358/22.  
 XX

DR N-PSDB; AAF60733.

XX Determining the effect of a test sample on UDP-glucose:glycoprotein  
PT glucosyltransferase (UGGT), useful for measuring UGGT activity,  
PT comprises exposing an acceptor substrate for UGGT to a labeled donor in  
PT the presence of UGGT -

XX Example 1; Fig 8; 95pp; English.

XX The present invention relates to a method for determining the effect of a  
CC test sample on UDP-glucose:glycoprotein glucosyltransferase (UGGT)  
CC activity. The method comprises exposing an acceptor substrate for UGGT to  
CC a labelled donor in the presence of the test sample and UGGT. The method  
CC is useful for determining UGGT activity. In particular, the method is  
CC useful in glucosyltransferase assay and kinetics measurement for  
CC determining UGGT activity. UGGT is a soluble enzyme of the endoplasmic  
CC reticulum which catalyses the addition of a glucose residue onto  
CC asparagine-linked oligosaccharides, which are present on incorrectly  
CC folded glycoproteins. The present sequence is the honeybee mellitin  
CC signal peptide. This sequence was used in the construction of an  
CC expression vector for rat UGGT (see AAF60732 and AAB72436).

XX Sequence 21 AA;

Query Match 94.4%; Score 102; DB 22; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.3e-09; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0;

QY 1 MKFLVNVALVFMVVIYSIYA 21  
Db 1 mkflvnvalvfmvviysiya 21  
|||||

RESULT 8

AAB36293  
ID AAB36293 standard; Protein; 49 AA.

XX AAB36293;

XX 23-FEB-2001 (first entry)

XX Human GIL-19/AE289 protein sequence FLAG epitope tag.

XX Human; GIL-19/AE289; IL-10; Interleukin-10; nutrition;  
KW cell proliferation; immune stimulation; immune suppression;  
KW haematopoiesis regulation; tissue growth; inflammation; cancer.

XX Synthetic.

XX WO200065027-A2.

XX 02-NOV-2000.

XX 28-APR-2000; 2000WO-US11479.

XX 28-APR-1999; 99US-0131473.

XX (GEMY ) GENETICS INST INC.

XX Jacobs K, Fouser L, Spaulding V, Xuan D;

XX WPI; 2000-687325/67.

XX Human GIL-19 protein that shows a high degree of homology to IL  
PT (interleukin)-10, useful in upregulation of humoral immune responses,  
PT as an antiinflammatory agent and as a modulator of immune responses  
PT associated with injury -

XX Example; Page 50; 60pp; English.

XX The present invention provides the protein and coding sequences for the  
CC novel human GIL-19/AE289 protein. The protein shows homology to

CC Interleukin-10 (IL-10) and is assumed to be a cytokine. It can be used  
CC in the regulation of cell proliferation and differentiation,  
CC haematopoiesis, immune stimulation or suppression, tissue growth and  
CC tumour inhibition. In addition, it also has uses in the treatment of  
CC inflammation and in nutrition.

XX Sequence 49 AA;

Query Match 94.4%; Score 102; DB 21; Length 49;  
Best Local Similarity 100.0%; Pred. No. 3.2e-09; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0;

QY 1 MKFLVNVALVFMVVIYSIYA 21  
Db 1 mkflvnvalvfmvviysiya 21  
|||||

RESULT 9

AAY97249

ID AAY97249 standard; Protein; 177 AA.

XX AAY97249;

XX 04-DEC-2000 (first entry)

XX Cytomegalovirus US2t-DL6 fusion protein.

XX US2; hCMV; major histocompatibility complex; MHC; class I; class II;  
KW antigen presentation; inhibition; CD8-positive; CD4-positive; T cell;  
KW transplant; gene therapy; immunosuppressive; fusion; DL6.

XX Chimeric - Human cytomegalovirus.

XX Chimeric - Apis sp.

XX Chimeric - Homo sapiens.

XX Chimeric - Synthetic.

XX Key Location/Qualifiers

XX 1..21

XX /label= Bee\_mellitin\_signal\_peptide

XX 22..157

XX /label= Soluble\_US2\_protein

XX /note= "Single glycine spacer"

XX 159..177

XX /label= DL6\_epitope

XX WO200046361-A1.

XX 10-AUG-2000.

XX 02-FEB-2000; 2000WO-US02740.

XX 02-FEB-1999; 99US-0118287.

XX (UYOR-) UNIV OREGON HEALTH SCI.

XX Johnson DC, Tomazin R, Boname J, Hegde NR;

XX WPI; 2000-506069/45.

XX N-PSDB; AAA53814.

XX Inhibiting recognition of cellular tissue by CD8+ and CD4+ T cells, to  
PT treat or prevent autoimmune diseases, and to improve gene therapy,  
PT comprises introducing human cytomegalovirus US2 protein into cells  
XX Example 12; Page 49; 53pp; English.

XX Human cytomegalovirus (hCMV) US2 protein, which has previously been shown  
CC to block the major histocompatibility complex (MHC) class I antigen  
CC presentation pathway, blocks the MHC class II pathway. The US2 protein  
CC does not have to be mutated to cause inhibition of the MHC class II  
CC pathway. The binding domain recognizes MHC I heavy chains, MHC II alpha

Qy	1	MKFLVNVALVFVVVWISYIYA	21
		:     :	
Db	1	mkflvdvalvfvmvviysiya	21
RESULT 11			
	AAW87499		
ID	AAW87499	standard; peptide; 21 AA.	
XX			
AC	AAW87499;		
XX			
DT	22-FEB-1999	(first entry)	
XX			

irion; transfection; cytopathic; honey bee; excretion; peptide;  
pis sp.

US5849523-A.  
15-DEC-1998.

07-JUN-1995; 95US-0485341.  
04-NOV-1994; 94US-0334669.

(Q RESE ) RESEARCH CORP TECHNOLOGIES INC.  
Afanasiev BN, Beaty BJ, Carlson JO, Higgins DR;  
Thibault KJ;

WPT: 1999-069725/06.

Recombinant protein production in mosquito cells - using Aedes

Column 6: 23pp: English:

The invention relates to the use of a *Aedes aegypti* densonucleosis virus (AeDNV) as a heterologous gene expression system. A recombinant protein using the vectors of the invention can be produced by a method that comprises (1) transfecting an insect cell that stably expresses AeDNV structural proteins VP1 and VP2 with an expression vector comprising the structural proteins VP1 and VP2 with an expression cassette flanked by 5' and 3' termini of AeDNV genomic DNA flanking an expression cassette coding for the protein.

(AeDNV) as a heterologous gene expression system. AeDNV vectors of the invention can be produced by a method that uses the vectors of the invention can be produced by a method that comprises (1) transfecting an insect cell that stably expresses AeDNV structural proteins VP1 and VP2 with an expression vector comprising the 5' and 3' termini of AeDNV genomic DNA flanking an expression cassette comprising an AeDNV promoter linked to a sequence coding for the protein of interest and with an expression vector comprising the AeDNV nonstructural proteins NS1 and NS2 encoding genes; (2) obtaining infectious AeDNV from the transfected cell; (3) infecting a second insect cell with the virions; culturing the second cell under conditions such that the protein is expressed; and (4) recovering the protein from the second cell. Aedes albopictus C6/36 cells can be transfected with AeDNV vectors without cytopathic effects. The present sequence represents a honey bee mellitin excretion peptide. This signal peptide can be used in the AeDNV system for the expression of nuclear localisation.

comprises (1) transfecting an insect cell that stably expresses structural proteins VP1 and VP2 with an expression vector comprising the 5' and 3' termini of AeDNV genomic DNA flanking an expression cassette and with an AeDNV promoter linked to a sequence coding for the protein and (2) obtaining infectious AeDNV nonstructural proteins NS1 and NS2 encoding genes; (3) infecting a second insect cell with the transfected cell; (4) recovering the protein from the second cell. Aedes albopictus C6/36 cells can be transfected with AeDNV vectors without cytopathic effects. The present sequence represents a honey bee mellitin excretion peptide. This signal peptide can be used in the AeDNV expression of nuclear localisation.

5' and 3' termini of AeDNV genomic DNA flanking an expression cassette comprising an AeDNV promoter linked to a sequence coding for the protein and with an expression vector comprising the AeDNV nonstructural proteins NS1 and NS2 encoding genes; (2) obtaining infectious virions from the transfected cell; (3) infecting a second insect cell with the virions; culturing the second cell under conditions such that the protein is expressed; and (4) recovering the protein from the second cell. Aedes albopictus C6/36 cells can be transfected with AeDNV vectors without cytopathic effects. The present sequence represents a honey bee mellitin excretion peptide. This signal peptide can be used in the AeDNV expression system for the localization of nuclear localisation.

comprising an expression vector comprising the AeDNV nonstructural protein NS1 and NS2 encoding genes; (2) obtaining infectious virions from the transfected cell; (3) infecting a second insect cell with the virions, culturing the second cell under conditions such that the protein is expressed, and (4) recovering the protein from the second cell. Aedes albopictus C6/36 cells can be transfected with AeDNV vectors without cytopathic effects. The present sequence represents a honey bee mellitin excretion peptide. This signal peptide can be used in the AeDNV

NS1 and NS2 enclosing genes; (2) infecting a second insect cell with the virions; (3) infecting a second cell under conditions such that the protein is transfected cell; (4) recovering the protein from the second cell. Aedes albopictus C6/36 cells can be transfected with AedNV vectors without cytopathic effects. The present sequence represents a honey bee mellitin exon 1A deletion construct. This signal peptide can be used in the AedNV for nuclear localisation.

culturing the second cell under conditions that induce AedNV expression, and (4) recovering the protein from the second cell. Aedes albopictus C6/36 cells can be transfected with AedNV vectors without cytotoxic effects. The present sequence represents a honey bee mellitin excretion peptide. This signal peptide can be used in the AedNV for nuclear localisation.

allopictus C6/36 cells can be transfected with AedDV vectors to elicit cytopathic effects. The present sequence represents a honey bee mellitin excretion peptide. This signal peptide can be used in the AedDV for transfection or nuclear localisation.

mellitin excretion peptide. This signal peptide can be used in the recombinant cytoplasmic expression system for the secretion or nuclear localisation.

Sequence	21 AA;
Very Match	89.8%;
St Local Similarity	95.2%;
tches	20; Conservative
Score	97; DB
Pred. NO.	7.9e-09;
Mismatches	1;
Indels	0;
Gaps	0;

```
1 mkf\vdva\vfmvvsviya 21
```

RESULT 12

AAW64613  
 ID AAW64613 standard; Protein: 26 AA.  
 XX  
 AC AAW64613;  
 XX  
 DT 03-NOV-1998 (first entry)  
 XX  
 DE pMelBac secretion signal peptide motif.  
 XX  
 KW Target; imaging agent; epithelium; transepithelial transport; diagnosis;  
 KW transcytosis; disease; basolateral; internalisation; J chain.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9830591-A1.  
 XX  
 PD 16-JUL-1998.  
 XX  
 PF 09-JAN-1998; 98WO-US00339.  
 XX  
 PR 10-JAN-1997; 97US-0782480.  
 XX  
 PA (EPIC-) EPICYTE PHARM INC.  
 XX  
 PI Fitchen JH, Hein MB, Hiatt AC;  
 XX  
 DR WPI; 1998-399066/34.  
 DR N-PSDB; AAV49671.  
 XX  
 PT New epithelial tissue targeting agent - used to deliver imaging  
 PT agents to an epithelial surface for internalisation; useful in  
 PT diagnosis  
 XX  
 PS Example 1c; Page 89; 118pp; English.  
 XX  
 CC This sequence represents a secretion signal from the plasmid pMelBac  
 CC which is used in a method involving the construction of a target molecule  
 CC from human J chain protein fragments. This construct is used in a method  
 CC to target imaging agents to epithelial surfaces at which they may remain  
 CC or undergo transepithelial transport via transcytosis. At least one  
 CC imaging agent is linked to the targeting molecule comprising a  
 CC polypeptide that (a) forms a closed covalent loop, (b) contains at least  
 CC 3, preferably 4, peptide domains having beta-sheet character separated by  
 CC domains lacking beta-sheet character and (c) is not full length dimeric  
 CC IGA. The imaging agents are useful in the diagnosis of disease. The  
 CC target molecule is also capable of specifically binding to a basolateral  
 CC factor associated with an epithelial surface to cause internalisation of  
 CC a biological agent linked to the target molecule.  
 XX  
 SQ Sequence 26 AA;  
 Query Match 86.6%; Score 93.5; DB 19; Length 26;  
 Best Local Similarity 95.5%; Pred. No. 3.5e-08;  
 Matches 21; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 MKFLVNVALFMVVISYIYAD 22  
 Db 1 mkflvnval-fmvvisyiyad 21  
 RESULT 13  
 AAW61590  
 ID AAW61590 standard; Protein: 26 AA.  
 XX  
 AC AAW61590;  
 XX  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE Targeting molecule secretion signal and multiple cloning site.  
 XX  
 KW J chain; targeting molecule; epithelial; beta-sheet; asthma;  
 KW cancer; inflammatory disorder; autoimmune disorder; celiac disease;

KW colitis; pneumonia; cystic fibrosis.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9830592-A1.  
 XX  
 PD 16-JUL-1998.  
 XX  
 PF 09-JAN-1998; 98WO-US00542.  
 XX  
 PR 10-JAN-1997; 97US-0782481.  
 XX  
 PA (EPIC-) EPICYTE PHARM INC.  
 XX  
 PI Fitchen JH, Hein MB, Hiatt AC;  
 XX  
 DR WPI; 1998-399067/34.  
 DR N-PSDB; AAV45274.  
 XX  
 PT New epithelial tissue targeting agent - used to deliver  
 PT biologically active compounds to an epithelial surface for  
 PT internalisation  
 XX  
 PS Example 1; Page 44; 142pp; English.  
 XX  
 CC The secretion signal and multiple cloning site are used in the assembly  
 CC of a targeting molecule (TM). The TMs are used to target biological  
 CC agents to epithelial surfaces at which they can be internalised. The TMs  
 CC comprise a polypeptide that: (a) forms a closed covalent loop; (b)  
 CC contains at least 3, preferably 4, peptide domains having beta-sheet  
 CC character separated by domains lacking beta-sheet character; and  
 CC (c) is not full length dimeric IGA. The TMs are useful to prevent and/or  
 CC treat diseases associated with epithelial surfaces, e.g. asthma, cancer,  
 CC (myco)bacterial, viral or fungal infection, inflammatory disorders,  
 CC autoimmune disorders, celiac disease, colitis, pneumonia and cystic  
 CC fibrosis.  
 XX  
 SQ Sequence 26 AA;  
 Query Match 86.6%; Score 93.5; DB 19; Length 26;  
 Best Local Similarity 95.5%; Pred. No. 3.5e-08;  
 Matches 21; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 MKFLVNVALFMVVISYIYAD 22  
 Db 1 mkflvnval-fmvvisyiyad 21  
 RESULT 14  
 AAW85766  
 ID AAW85766 standard; Peptide: 26 AA.  
 XX  
 AC AAW85766;  
 XX  
 DT 27-SEP-1999 (first entry)  
 XX  
 DE Secretion signal and multiple cloning site.  
 XX  
 KW Targeting molecule; J chain; immunoglobulin; IgM; IgA; substrate;  
 KW epithelial cell; cancer; treatment; therapy;  
 KW non-small cell lung carcinoma; breast carcinoma; colon carcinoma;  
 KW ovarian carcinoma; prostate carcinoma; endometriosis;  
 KW viral infection; inflammation.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9920310-A1.  
 XX  
 PD 29-APR-1999.  
 XX  
 PF 20-OCT-1998; 98WO-US22304.  
 XX

Mon Feb 4 15:23:46 2002

XX Barbet AF, Burridge MJ, Ganta RR, Mahan SM, McGuire TC;  
 PI Nyika A, Rurangirwa FR;  
 XX WPI; 1998-251232/22.  
 DR N-PSDB; AAV07181.  
 XX Composition containing nucleic acid encoding rickettsial antigen  
 PT useful for, e.g. stimulating protective immune response in humans or  
 PT animals  
 XX Claim 3; Fig 3A; 39pp; English.  
 PS This polypeptide comprises the major antigen protein 2 gene (MAP2)  
 CC of Ehrlichia canis. It is encoded by the MAP2 gene (see AAV07181).  
 CC A claimed composition comprises a nucleic acid (see AAV07176-82)  
 CC encoding a polypeptide (see AAW51088-99) that elicits a protective  
 CC immune response against a rickettsial pathogen. The nucleic acid  
 CC is used, in human or veterinary medicine, in vaccines to protect  
 CC against Rickettsia, Ehrlichia, Anaplasma and Cowdria species. The  
 CC nucleic acid does not replicate in the host but remains episomal  
 CC and capable of expressing polypeptide for at least 19 mth. The  
 CC Ehrlichia antigenic polypeptides can also be used diagnostically to  
 CC detect antibodies associated with Ehrlichia infection (claimed).  
 XX Sequence 205 AA;  
 SQ

Query Match 50.9%; Score 55; DB 19; Length 205;  
 Best Local Similarity 40.0%; Pred. No. 0.31;  
 Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKFLVNVVLFVWVYISYIY 20  
 :||:|||||:|:|:|:|  
 Db 4 lmkflvnval-fmvvysiyad 23

Search completed: January 29, 2002, 10:59:20  
 Job time: 2244 sec

PR 20-OCT-1997; 97US-0954211.  
 XX (EPIC-) EPICYTE PHARM INC.  
 PA Fitchen JH, Hein MB, Hiatt AC;  
 XX WPI; 1999-288174/24.  
 DR N-PSDB; AAX08824.  
 XX Targeting molecule useful in drug delivery for treating cancer,  
 PT viral infection or inflammatory disorders  
 PT Example 1; Page 43; 102pp; English.  
 PS A targeting agent for improving the delivery of drugs to target  
 XX cells, particularly for delivery of enzymes, binding agents,  
 CC inhibitors, nucleic acids, carbohydrates and lipids, is new.  
 CC The targeting agent comprises a polypeptide which forms a closed  
 CC covalent loop and contains at least three peptide domains having  
 CC beta-sheet character, each of the domains being separated by domains  
 CC lacking beta-sheet character. The targeting molecule preferably  
 CC comprises all or a portion of a native J chain sequence. J chain  
 CC is a 15 kD protein that, in vivo, links IgM or IgA monomers to  
 CC form pentameric IgM or dimeric IgA. The polypeptide is linked to  
 CC at least one biological agent which is capable of entering and  
 CC killing a non-polarised epithelial cell. The targeting molecule may  
 CC be linked to the biological agent by a substrate for an  
 CC intracellular or extracellular enzyme which is associated with or  
 CC secreted by the non-polarised target cell. The targeting molecule  
 CC can be used in a pharmaceutical composition for treating a patient  
 CC afflicted with a disease associated with non-polarised epithelial  
 CC cells, especially cancer e.g. non-small cell lung carcinoma,  
 CC breast carcinoma, colon carcinoma, ovarian carcinoma, prostate  
 CC carcinoma and endometrios, viral infection or inflammatory  
 CC disorders. This sequence is a secretion signal. The nucleotide  
 CC encoding this sequence is used in an expression vector used for the  
 CC construction of target molecules described.  
 XX Sequence 26 AA;  
 SQ

Query Match 86.6%; Score 93.5; DB 20; Length 26;  
 Best Local Similarity 95.5%; Pred. No. 3.5e-08;  
 Matches 21; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKFLVNVVLFVWVYISYIYAD 22  
 ||||| ||||| ||||| |||||  
 Db 1 mkflvnval-fmvvysiyad 21

RESULT 15  
 AAW51098  
 ID AAW51098 standard; Protein; 205 AA.  
 XX AC AAW51098;  
 XX 14-SEP-1998 (first entry)  
 XX Ehrlichia canis major antigenic protein 2 (MAP2).  
 XX MAP2; major antigenic protein 2; rickettsia; DNA vaccine.  
 XX Ehrlichia canis.  
 XX WO9816554-A1.  
 XX 23-APR-1998.  
 XX 17-OCT-1997; 97WO-US19044.  
 XX 17-OCT-1996; 96US-0733230.  
 XX (UYFL ) UNIV FLORIDA.





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:16:04 ; Search time 1760.55 Seconds  
(without alignments)  
55.199 Million cell updates/sec

Title: US-09-763-397A-2

Perfect score: 1923

Sequence: 1 MKELVNVALFVWVYISYIY.....DFFGISYKVLAKYKDDLE 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main :

1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	1923	100.0	350	21	US-09-763-397A-2
2	215.5	11.2	412	3	US-07-677-539B-6
3	215.5	11.2	412	3	US-07-727-636-6
4	215.5	11.2	412	13	US-08-948-885-18
5	210.5	10.9	423	13	US-08-932-929A-1
6	210.5	10.9	424	3	US-07-842-694-2
7	210	10.9	127	15	US-09-125-031A-8
8	210	10.9	127	15	US-09-125-031A-8
9	210	10.9	127	15	US-09-134-333-8

10	208.5	10.8	388	3	US-07-677-539B-5	Sequence 5, Appli
11	208.5	10.8	388	3	US-07-727-636-5	Sequence 5, Appli
12	207.5	10.8	396	22	US-09-820-843A-31	Sequence 31, Appli
13	207.5	10.8	424	11	US-08-760-797-3	Sequence 3, Appli
14	207.5	10.8	424	13	US-08-903-084-3	Sequence 3, Appli
15	207.5	10.8	424	13	US-08-932-929-3	Sequence 3, Appli
16	207.5	10.8	424	13	US-08-932-929A-3	Sequence 3, Appli
17	206.5	10.7	424	11	US-08-760-797-1	Sequence 1, Appli
18	206.5	10.7	424	13	US-08-903-084-1	Sequence 1, Appli
19	206.5	10.7	424	13	US-08-932-929-1	Sequence 1, Appli
20	203.5	10.6	402	19	US-09-500-376-16	Sequence 16, Appli
21	200.5	10.4	394	3	US-07-867-768A-2	Sequence 2, Appli
22	200.5	10.4	394	5	US-08-195-705-2	Sequence 2, Appli
23	200.5	10.4	394	19	US-09-500-376-2	Sequence 2, Appli
24	198.5	10.3	95	15	US-09-125-031-2	Sequence 2, Appli
25	198.5	10.3	95	15	US-09-125-031A-2	Sequence 2, Appli
26	198.5	10.3	95	15	US-09-125-031B-2	Sequence 2, Appli
27	198.5	10.3	95	15	US-09-134-333-2	Sequence 2, Appli
28	198.5	10.3	108	15	US-09-125-031-10	Sequence 10, Appli
29	198.5	10.3	108	15	US-09-125-031A-10	Sequence 10, Appli
30	198.5	10.3	108	15	US-09-125-031B-10	Sequence 10, Appli
31	198.5	10.3	108	15	US-09-134-333-10	Sequence 10, Appli
32	198.5	10.3	116	15	US-09-125-031-5	Sequence 5, Appli
33	198.5	10.3	116	15	US-09-125-031A-5	Sequence 5, Appli
34	198.5	10.3	116	15	US-09-134-333-5	Sequence 5, Appli
35	198.5	10.3	116	15	US-09-125-031B-5	Sequence 5, Appli
36	198.5	10.3	127	15	US-09-125-031B-8	Sequence 8, Appli
37	196.5	10.2	96	17	US-09-311-817-1	Sequence 1, Appli
38	196.5	10.2	355	15	US-09-175-683-11	Sequence 11, Appli
39	196.5	10.2	355	15	US-09-175-683B-9	Sequence 9, Appli
40	196.5	10.2	361	15	US-09-175-683-12	Sequence 12, Appli
41	196.5	10.2	361	15	US-09-175-683B-10	Sequence 10, Appli
42	196.5	10.2	376	17	US-09-311-817-2	Sequence 2, Appli
43	196.5	10.2	384	19	US-09-500-376-8	Sequence 8, Appli
44	196.5	10.2	594	9	US-08-593-006-2	Sequence 2, Appli
45	196.5	10.2	594	15	US-09-117-415-2	Sequence 2, Appli

## ALIGNMENTS

RESULT .1

US-09-763-397A-2  
; Sequence 2, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Recombinant DNA/Protein  
US-09-763-397A-2

Query Match 100.0%; Score 1923; DB 21; Length 350;  
Best Local Similarity 100.0%; Pred. No. 1.5e-164;  
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLVNVALVPMVYIYYIADHHHHHHKHLKOPGDNPNWSPSCSVTCGPKDELDYEN 60  
DB 1 MKFLVNVALVPMVYIYYIADHHHHHHKHLKOPGDNPNWSPSCSVTCGPKDELDYEN 60  
QY 61 DIEKKICKMEKCSVFNWVNSGCFRHLDERECKCILEDGSGNGKKTCTCKPDSKP 120  
DB 61 DIEKKICKMEKCSVFNWVNSGCFRHLDERECKCILEDGSGNGKKTCTCKPDSKP 120  
QY 121 IVQYDNFNANPNANPNAGNEDIPHYNEFSAIDLGNAEKYDKMDEPOHYGKSLTPLE 180  
DB 121 IVQYDNFNANPNANPNAGNEDIPHYNEFSAIDLGNAEKYDKMDEPOHYGKSLTPLE 180  
QY 181 ELYKPNDSLVQYIKANSKFIGITELSNFTFINNAGOHGMHGNREDERTLTKEYEDIVL 240  
DB 181 ELYKPNDSLVQYIKANSKFIGITELSNFTFINNAGOHGMHGNREDERTLTKEYEDIVL 240  
QY 241 KEFTYMINFGRGQNYWEHPYQKSDQPKQYEQHLTDYERIKEGKPLDRFGNIYDHYHSS 300  
DB 241 KEFTYMINFGRGQNYWEHPYQKSDQPKQYEQHLTDYERIKEGKPLDRFGNIYDHYHSS 300  
QY 301 PSTKSSPSNVKASLASLATRLMKKFAEIRDFPGISYVEKVLAKYKDDLE 350  
DB 301 PSTKSSPSNVKASLASLATRLMKKFAEIRDFPGISYVEKVLAKYKDDLE 350

## RESULT 2

US-07-677-539B-6  
; Sequence 6, Application US/07677539B  
; GENERAL INFORMATION:  
; APPLICANT: LAL, ALTAF A.  
; TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM  
; TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L STREET, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/677-539B  
; FILING DATE: 19911205  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SCOTT, WATSON T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: 5683/91540  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3067  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 412 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-07-677-539B-6

Query Match 11.2%; Score 215.5; DB 3; Length 412;  
Best Local Similarity 62.2%; Pred. No. 2.8e-10;  
Matches 46; Conservative 2; Mismatches 7; Indels 19; Gaps 3;

QY 29 KH-----KKLKQPGDGNPNWSPSCSVTCG-----KPKDELDYENDIEKKICKME 70  
DB 329 KHIEQYLLKTKN-SISTEWSPCSVTCGNGIQVRIKPGSANKPKDELDYENDIEKKICKME 387  
QY 71 KCSSVFNWVNSG 84  
DB 388 KCSSVFNWVNSG 401  
RESULT 3  
US-07-727-636-6  
; Sequence 6, Application US/07727636  
; GENERAL INFORMATION:  
; APPLICANT: LAL, ALTAF A.  
; APPLICANT: GOLDMAN, IRA F.  
; TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM  
; TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L STREET, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/727-636  
; FILING DATE: 19910710  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SCOTT, WATSON T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: 5683/91540  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3067  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 412 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-07-727-636-6

Query Match 11.2%; Score 215.5; DB 3; Length 412;  
Best Local Similarity 62.2%; Pred. No. 2.8e-10;  
Matches 46; Conservative 2; Mismatches 7; Indels 19; Gaps 3;

QY 29 KH-----KKLKQPGDGNPNWSPSCSVTCG-----KPKDELDYENDIEKKICKME 70  
DB 329 KHIEQYLLKTKN-SISTEWSPCSVTCGNGIQVRIKPGSANKPKDELDYENDIEKKICKME 387  
QY 71 KCSSVFNWVNSG 84  
DB 388 KCSSVFNWVNSG 401

## RESULT 4

US-08-948-885-18  
; Sequence 18, Application US/08948885  
; GENERAL INFORMATION:  
; APPLICANT: Jessell, Thomas M. and Avihu Klar  
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A  
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SFONDIN  
; NUMBER OF SEQUENCES: 20

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Cooper & Dunham LLP  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/948,885  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/313,288  
;; FILING DATE: January 5, 1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: White, John P.  
;; REGISTRATION NUMBER: 28,678  
;; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 278-0400  
;; TELEFAX: (212) 391-0526  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 18:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 412 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-948-885-18

Query Match 11.2%; Score 215.5; DB 13; Length 412;  
Best Local Similarity 62.2%; Pred. No. 2.8e-10;  
Matches 46; Conservative 2; Mismatches 7; Indels 19; Gaps 3;

Qy 29 KH-----KKLKPGDGNPSCVTCG-----KPKDELVDYENDIEKKICKME 70  
Db 329 KHIEQYLKKIKN-SISTEMSPSCVTCGNGIQVRIKPGSANKPKDELVDYENDIEKKICKME 387  
Qy 71 KCSSVFNVSNSG 84  
Db 388 KCSSVFNVSNSIG 401

RESULT 5  
US-08-932-929A-1  
;; Sequence 1, Application US/08932929A  
;; GENERAL INFORMATION:  
;; APPLICANT: Cohen, Joseph  
;; TITLE OF INVENTION: Hybrid Protein Between CS  
;; TITLE OF INVENTION: from Plasmodium and HBSAG  
;; NUMBER OF SEQUENCES: 4  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SmithKline Beecham Corporation  
;; STREET: 709 Swedeland Road  
;; CITY: King of Prussia  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19406  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/932,929A

;; FILING DATE: 18-SEPT-1997  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/663,371  
;; FILING DATE: 13-JUNE-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Baumeister, Kirk  
;; REGISTRATION NUMBER: 33,833  
;; REFERENCE/DOCKET NUMBER: B45015-1FWC2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 610-270-5096  
;; TELEFAX: 610-270-5090  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 423 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
US-08-932-929A-1

Query Match 10.9%; Score 210.5; DB 13; Length 423;  
Best Local Similarity 63.4%; Pred. No. 8.4e-10;  
Matches 45; Conservative 1; Mismatches 6; Indels 19; Gaps 3;

Qy 29 KH-----KKLKPGDGNPSCVTCG-----KPKDELVDYENDIEKKICKME 70  
Db 123 KHIEQYLKKIKN-SISTEMSPSCVTCGNGIQVRIKPGSANKPKDELVDYENDIEKKICKME 181  
Qy 71 KCSSVFNVSNS 81  
Db 182 KCSSVFNVSNS 192

RESULT 6  
US-07-842-694-2  
;; Sequence 2, Application US/07842694  
;; GENERAL INFORMATION:  
;; APPLICANT: Cohen, Joseph  
;; TITLE OF INVENTION: Malaria Vaccine Antigen  
;; NUMBER OF SEQUENCES: 2  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SmithKline Beecham Corporation  
;; STREET: P.O. Box 1539 / Corporate Patents  
;; CITY: King of Prussia  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19406-0939  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/842,694  
;; FILING DATE: 19920227  
;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Canter, Carol G.  
;; REGISTRATION NUMBER: 31151  
;; REFERENCE/DOCKET NUMBER: B45015  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 424 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
US-07-842-694-2



```
RESULT 10
US-07-677-539B-5
; Sequence 5, Application US/07677539B
; GENERAL INFORMATION:
; APPLICANT: LAL, ALTAFA A.
; APPLICANT: GOLDMAN, IRA F.
; TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM
; TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/677.539B
; FILING DATE: 19911205
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: 5683/91540
; TELEPHONE: 202-861-3067
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-677-539B-5

Query Match 10.8%; Score 208.5; DB 3; Length 388;
Best Local Similarity 58.3%; Pred. No. 1.1e-09;
Matches 42; Conservative 4; Mismatches 13; Indels 13; Gaps 1;

Qy 26 HHHKHKLLKQPGDGNPSPCVTC-----GPKDLDYENDIEKKICKMEKC 72
Db 306 HIEFLKQIQNNLSTWSPCVTCGNGIQVRIKPGSAGKPKDQLDYENDLEKKICKMEKC 365

Qy 73 SSVFNVNNSG 84
Db 366 SSVFNVNNSG 377

RESULT 11
US-07-727-636-5
; Sequence 5, Application US/07727636
; GENERAL INFORMATION:
; APPLICANT: LAL, ALTAFA A.
; APPLICANT: GOLDMAN, IRA F.
; TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM
; TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/727.636
; FILING DATE: 19910710
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: 5683/91540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3067
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-727-636-5

Query Match 10.8%; Score 208.5; DB 3; Length 388;
Best Local Similarity 58.3%; Pred. No. 1.1e-09;
Matches 42; Conservative 4; Mismatches 13; Indels 13; Gaps 1;

Qy 26 HHHKHKLLKQPGDGNPSPCVTC-----GPKDLDYENDIEKKICKMEKC 72
Db 306 HIEFLKQIQNNLSTWSPCVTCGNGIQVRIKPGSAGKPKDQLDYENDLEKKICKMEKC 365

Qy 73 SSVFNVNNSG 84
Db 366 SSVFNVNNSG 377

RESULT 12
US-09-820-843A-31
; Sequence 31, Application US/09820843A
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PRO
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Circumsporozoite (CS) protein
; NAME/KEY: misc_feature
; OTHER INFORMATION: gi|4493889
; US-09-820-843A-31

Query Match 10.8%; Score 207.5; DB 22; Length 396;
Best Local Similarity 51.8%; Pred. No. 1.4e-09;
Matches 43; Conservative 4; Mismatches 9; Indels 27; Gaps 2;

Qy 29 KHKLLKQPGD-----GNPSPSPCVTCG-----KPKDLDYEND 61
Db 303 KNNNEEPSPDKHIKYLKQNLSTWSPCVTCGNGIQVRIKPGSANKPKDLDYAND 362

Qy 62 IEKKICKMEKCSSVFNVNNSG 84
```





```

;
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,929
; FILING DATE: 18-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/663,371
; FILING DATE:
; APPLICATION NUMBER: US 08/244,085
; FILING DATE: 21-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610 270-5096
; TELEFAX: 610 270-5090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-932-929-3

```

```

Query Match      10.8%; Score 207.5; DB 13; Length 424;
Best Local Similarity 51.8%; Pred No. 1.6e-09;
Matches 43; Conservative 4; Mismatches 9; Indels 27; Gaps 2;

QY 29 KHKKLKQPGD-----GNWSPCSVTG-----KPKDELDYND 61
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 110 KNNNEEPSDKHIKEYLNKIQNSLSTWSPCSVTGCGIQVRIKPGSANKPKDELDYAND 169

QY 62 IEKKICKMEKCSVFNVNNSG 84
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 170 IEKKICKMEKCSVFNVNNSIG 192

```

Search completed: January 29, 2002, 10:55:58  
Job time: 2394 sec

**THIS PAGE BLANK (uspto)**



;; FILING DATE: 14-AUG-1997  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/915,503  
;; FILING DATE: 14-AUG-1997  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: WO PCT/US97/17618  
;; FILING DATE: 01-OCT-1997  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: WO PCT/US97/17885  
;; FILING DATE: 01-OCT-1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Apple, Randolph Ted  
;; REGISTRATION NUMBER: 36,429  
;; REFERENCE/DOCKET NUMBER: 015389-002610US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 576-0200  
;; TELEFAX: (415) 576-0300  
;; INFORMATION FOR SEQ ID NO: 613:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FEATURE:  
;; NAME/KEY: Protein  
;; LOCATION: 1..1189  
;; OTHER INFORMATION: /note= "fusion protein composed of  
;; OTHER INFORMATION: melittin signal sequence and full length  
;; OTHER INFORMATION: hTMT protein"  
US-08-974-549A-613

Query Match 100.0%; Score 108; DB 4; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 8.3e-09;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLVNVALVFMVYISYIYAD 22  
DB 1 MKFLVNVALVFMVYISYIYAD 22

RESULT 2  
US-08-960-190A-36  
; Sequence 36, Application US/08960190A  
; Patent No. 6232445  
; GENERAL INFORMATION:  
; APPLICANT: Rhode, Peter R.  
; APPLICANT: Acevedo, Jorge  
; APPLICANT: Burkhardt, Martin  
; APPLICANT: Jiao, Jin-an  
; APPLICANT: Wong, Hing C.  
; TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/960,190A  
; FILING DATE: 29-OCT-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:

;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Corless, Peter F.  
;; REGISTRATION NUMBER: 33,860  
;; REFERENCE/DOCKET NUMBER: 48002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-523-3400  
;; TELEFAX: 617-523-6440  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 36:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 21 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-960-190A-36  
Query Match 94.4%; Score 102; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1e-09;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKFLVNVALVFMVYISYIYA 21  
DB 1 MKFLVNVALVFMVYISYIYA 21  
RESULT 3  
US-08-334-669-2  
; Sequence 2, Application US/08334669  
; Patent No. 5627048  
; GENERAL INFORMATION:  
; APPLICANT: Higgins, David R.  
; APPLICANT: Thibault, Kelly J.  
; APPLICANT: Afanasiev, Boris N.  
; APPLICANT: Carlson, Jonathan O.  
; APPLICANT: Beatty, Barry  
; TITLE OF INVENTION: AEDES AEGYPTI DENSOVIRUS EXPRESSION  
; TITLE OF INVENTION: SYSTEM  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/334,669  
; FILING DATE: 04-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9485  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-334-669-2

```
Query Match      89.8%; Score 97; DB 1; Length 21;
Best Local Similarity 95.2%; Pred. No. 6e-09;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLVNVALFMVVIYSIYA 21
   |||||:|||||:|||||:
Db 1 MKFLVDVALFMVVIYSIYA 21

RESULT 4
US-08-485-341A-2
; Sequence 2, Application US/08485341A
; Patent No. 5849523
; GENERAL INFORMATION:
; APPLICANT: Afanasiev, Boris N.
; APPLICANT: Carlson, Jonathan O.
; APPLICANT: Beatty, Barry J.
; APPLICANT: Higgins, David R.
; APPLICANT: Thibault, Kelly J.
; TITLE OF INVENTION: AIDES AEGYPTI DENSOVIRUS
; TITLE OF INVENTION: EXPRESSION SYSTEM
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,341A
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 94852-I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)742-4343
; TELEFAX: (516)742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-485-341A-2

Query Match      89.8%; Score 97; DB 2; Length 21;
Best Local Similarity 95.2%; Pred. No. 6e-09;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLVNVALFMVVIYSIYA 21
   |||||:|||||:|||||:
Db 1 MKFLVDVALFMVVIYSIYA 21

RESULT 5
US-08-782-480-43
; Sequence 43, Application US/08782480
; Patent No. 6045774
; GENERAL INFORMATION:
; APPLICANT: Hiatt, Andrew C.
; APPLICANT: Hein, Mich B.
; APPLICANT: Hiatt, Andrew C.
; TITLE OF INVENTION: NOVEL EPITHELIAL CELL TARGETING AGENT
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,211
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; APPLICANT: Hein, Mich B.
; APPLICANT: Hiatt, Andrew C.
; TITLE OF INVENTION: NOVEL EPITHELIAL CELL TARGETING AGENT
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,211
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
```

```
Query Match      86.6%; Score 93.5; DB 3; Length 26;
Best Local Similarity 95.5%; Pred. No. 2.6e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY 1 MKFLVNVALFMVVIYSIYAD 22
   |||||:|||||:|||||:
Db 1 MKFLVNVALFMVVIYSIYAD 21
```

```
RESULT 6
US-08-954-211-43
; Sequence 43, Application US/08954211
; Patent No. 6251392
; GENERAL INFORMATION:
; APPLICANT: Hein, Mich B.
; APPLICANT: Hiatt, Andrew C.
; APPLICANT: Fitcher, John H.
; TITLE OF INVENTION: NOVEL EPITHELIAL CELL TARGETING AGENT
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,211
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
```

REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 310098.403  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-954-211-43

Query Match 86.6%; Score 93.5; DB 4; Length 26;  
Best Local Similarity 95.5%; Pred. No. 2.6e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKFLNVNVALFMVYISYIYAD 22  
Db 1 MKFLNVNVALFMVYISYIYAD 21

RESULT 7  
US-08-953-326-23  
; Sequence 23, Application US/08953326  
; Patent No. 6251872  
; GENERAL INFORMATION:  
; APPLICANT: Barbet, Anthony F.  
; APPLICANT: Ganta, Roman R.  
; APPLICANT: McGuire, Travis C.  
; APPLICANT: Burridge, Michael J.  
; APPLICANT: Nyika, Aceme  
; APPLICANT: Rurangirwa, Fred R.  
; APPLICANT: Mahan, Suman M.  
; TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of  
; TITLE OF INVENTION: Animals and Humans  
; FILE REFERENCE: UP-167C1  
; CURRENT APPLICATION NUMBER: US/08/953.326  
; EARLIER FILING DATE: 1997-10-17  
; EARLIER APPLICATION NUMBER: 08/953.326  
; EARLIER FILING DATE: 1997-10-17  
; EARLIER APPLICATION NUMBER: 08/733.230  
; EARLIER FILING DATE: 1996-10-17  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 205  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
US-08-953-326-23

Query Match 50.9%; Score 55; DB 4; Length 205;  
Best Local Similarity 40.0%; Pred. No. 0.18;  
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKFLNVNVALFMVYISYIY 20  
Db 4 IKFILNCLLFAAIFLGYSY 23

RESULT 8  
US-08-953-326-24  
; Sequence 24, Application US/08953326  
; Patent No. 6251872  
; GENERAL INFORMATION:  
; APPLICANT: Barbet, Anthony F.  
; APPLICANT: Ganta, Roman R.  
; APPLICANT: McGuire, Travis C.  
; APPLICANT: Burridge, Michael J.  
; APPLICANT: Nyika, Aceme  
; APPLICANT: Rurangirwa, Fred R.

APPLICANT: Mahan, Suman M.  
TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of  
TITLE OF INVENTION: Animals and Humans  
FILE REFERENCE: UP-167C1  
CURRENT APPLICATION NUMBER: US/08/953.326  
EARLIER FILING DATE: 1997-10-17  
EARLIER APPLICATION NUMBER: 08/953.326  
EARLIER FILING DATE: 1997-10-17  
EARLIER APPLICATION NUMBER: 08/733.230  
EARLIER FILING DATE: 1996-10-17  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 24  
LENGTH: 205  
TYPE: PRT  
ORGANISM: Ehrlichia chaffeensis  
US-08-953-326-24

Query Match 50.0%; Score 54; DB 4; Length 205;  
Best Local Similarity 35.0%; Pred. No. 0.26;  
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKFLNVNVALFMVYISYIY 20  
Db 4 IKFILNCLLFAAIFLGYSY 23

RESULT 9  
US-09-248-588-11  
; Sequence 11, Application US/09248588  
; Patent No. 6231864  
; GENERAL INFORMATION:  
; APPLICANT: Birkett, Ashley J.  
; TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and  
; TITLE OF INVENTION: their Derivatives  
; FILE REFERENCE: SYN-101.4564/69529  
; CURRENT APPLICATION NUMBER: US/09/248.588  
; CURRENT FILING DATE: 1999-02-11  
; EARLIER APPLICATION NUMBER: 60/074537  
; EARLIER FILING DATE: 1998-02-12  
; NUMBER OF SEQ ID NOS: 113  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 305  
; TYPE: PRT  
; ORGANISM: Hepatitis B virus  
US-09-248-588-11

Query Match 45.4%; Score 49; DB 4; Length 305;  
Best Local Similarity 55.0%; Pred. No. 2.3;  
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 FLVNVNVALFMVYISYIYAD 22  
Db 26 FLVTVPLVCTIVYDCLYMD 45

RESULT 10  
US-08-978-741-6  
; Sequence 6, Application US/08978741  
; Patent No. 6100076  
; GENERAL INFORMATION:  
; APPLICANT: Yang Wang, Michael W. Spellman  
; TITLE OF INVENTION: O-Fucosyltransferase  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA

ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/978,741  
FILING DATE: 26-No. 6100076-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/792498  
FILING DATE: 31  
ATTORNEY/AGENT INFORMATION:  
NAME: Svoboda, Craig G.  
REGISTRATION NUMBER: 39,044  
REFERENCE/DOCKET NUMBER: P1041P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1489  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-978-741-6

Query Match 45.4%; Score 49; DB 3; Length 397;  
Best Local Similarity 50.0%; Pred. No. 3;  
Matches 10; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 KFLVNVALFMVYISYIYA 21  
DB 3 KFLVNVALLLLLLLSGAWA 22

RESULT 11  
US-09-333-729A-7  
Sequence 7, Application US/09333729A  
Patent No. 6270987  
GENERAL INFORMATION:  
APPLICANT: Wang, Yang  
TITLE OF INVENTION: O-Fucosyltransferase  
FILE REFERENCE: P1041P1D1-Substitute  
CURRENT APPLICATION NUMBER: US/09/333,729A  
CURRENT FILING DATE: 1999-06-15  
PRIOR APPLICATION NUMBER: US 08/798,741  
PRIOR FILING DATE: 1997-11-26  
NUMBER OF SEQ ID NOS: 21  
SEQ ID NO 7  
LENGTH: 397  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Plasmid insert encoded protein.  
US-09-333-729A-7

Query Match 45.4%; Score 49; DB 4; Length 397;  
Best Local Similarity 50.0%; Pred. No. 3;  
Matches 10; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 KFLVNVALFMVYISYIYA 21  
DB 3 KFLVNVALLLLLLLSGAWA 22

RESULT 12  
US-08-687-289A-5  
Sequence 5, Application US/08687289A  
Patent No. 5981195

GENERAL INFORMATION:  
APPLICANT: Fuller, Forrest H.  
APPLICANT: Krapcho, Karen J.  
APPLICANT: Hammerland, Lance G.  
TITLE OF INVENTION: CHIMERIC RECEPTORS AND METHODS FOR  
IDENTIFYING COMPOUNDS ACTIVE AT  
TITLE OF INVENTION: METABOTROPIC GLUTAMATE RECEPTORS AND  
THE USE OF SUCH COMPOUNDS IN THE  
TITLE OF INVENTION: TREATMENT OF NEUROLOGICAL DISORDERS  
AND DISEASES  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq for Windows 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,289A  
FILING DATE: July 25, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/001,526  
FILING DATE: July 26, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 220/004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1058 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-687-289A-5

Query Match 41.7%; Score 45; DB 2; Length 1058;  
Best Local Similarity 45.0%; Pred. No. 35;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 KFLVNVALFMVYISYIYA 21  
DB 785 KFITFSMLFFIVWISFIPA 804

RESULT 13  
US-09-134-513-2  
Sequence 2, Application US/09134513  
Patent No. 6210964  
GENERAL INFORMATION:  
APPLICANT: Brown, Edward M.  
APPLICANT: Diaz, Ruben  
APPLICANT: Bai, Mei  
APPLICANT: Quinn, Stephen J.  
TITLE OF INVENTION: The Avian Extracellular Calcium-Sensing  
RECEPTOR  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Vinson & Elkins L.L.P.





```

: CLASSIFICATION: 435
: PRIOR APPLICATION DATA: Including application
: PRIOR APPLICATION DATA: described below: 9
: APPLICATION NUMBER: 08/353,784
: FILING DATE: 9 December, 1994
: APPLICATION NUMBER: PCT/US/94/12117
: FILING DATE: 21 October, 1994
: APPLICATION NUMBER: U.S. 08/292,827
: FILING DATE: 23 August, 1994
: APPLICATION NUMBER: U.S. 08/141,248
: FILING DATE: 22 October, 1993
: APPLICATION NUMBER: U.S. 08/009,389
: FILING DATE: 23 February, 1993
: APPLICATION NUMBER: U.S. 08/017,127
: FILING DATE: 12 February, 1993
: APPLICATION NUMBER: U.S. 07/934,161
: FILING DATE: 21 August, 1992
: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Heber, Sheldon O.
: REGISTRATION NUMBER: 38,179
: REFERENCE/DOCKET NUMBER: 213/006
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1078 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-484-565-7

```

```

Query Match 41.7% Score 45; DB 1; Length 1078;
Best Local Similarity 45.0%; Pred. No. 36;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Oy 2 KFLVNVALVFMVYISYIYA 21
||: ||: ||: ||: ||: ||:
Db 805 KFTFSMLIFFIVWISFIPA 824

```

Search completed: January 29, 2002, 10:59:59  
Job time: 2145 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: January 29, 2002, 10:14:44 ; Search time 310.82 Seconds  
(without alignments)  
83.410 Million cell updates/sec

Title: US-09-763-397A-2  
Perfect score: 1293  
Sequence: 1 MKFLNVALVEMVVISYIY.....DPFGISYERVLAKYKDDLE 350

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues  
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1923	100.0	350	21	Recombinant vaccine
2	352	13.1	309	12	NS1_81-RLfDelta9.
3	247	12.8	319	11	NS18RLfAuth plasm
4	247	12.8	319	12	NS1_81-RLfAuth. P
5	247	12.8	335	12	NS1_81 (NANP)4RLfau
6	247	12.8	335	12	NS1_81 (NVDP)4RLfau
7	243	12.6	327	12	NS1_81-RLfAuth + (
8	222	11.5	108	22	Merozoite surface
9	215.5	11.2	411	9	Sequence encoded b
10	215.5	11.2	412	7	CS protein of mala
11	211.5	11.0	412	9	Sequence encoded b

12	210.5	10.9	424	14	AAR37796	RTS protein. Synt
13	207.5	10.8	424	14	AAR37797	RTS* protein. Syn
14	198.5	10.3	116	18	AAW36103	PFMSPI(p19)A prote
15	198.5	10.3	116	18	AAW22592	PFMSPI(p19)A prote
16	198.5	10.3	127	18	AAW22593	PFMSPI(p19)S prote
17	198.5	10.3	127	18	AAW36102	PFMSPI(p19)S prote
18	196.5	10.2	96	22	AAW37608	Merozoite surface
19	196.5	10.2	355	20	AAV09372	Merozoite surface
20	196.5	10.2	355	20	AAV05832	Merozoite surface
21	196.5	10.2	361	20	AAV05833	Merozoite surface
22	196.5	10.2	361	20	AAV05833	Merozoite surface
23	195.5	10.2	1639	19	AAW54145	P. falciparum synt
24	195.5	10.2	376	20	AAV05834	Modified merozoite
25	193.5	10.2	376	20	AAV05834	Modified merozoite
26	193.5	10.1	180	11	AAV07290	Circumsporozoite a
27	187.5	9.8	375	22	AAW83926	A major merozoite
28	186	9.7	1654	6	AAV50777	Sequence of the P1
29	179	9.3	622	10	AAV91632	Rhoptry membrane a
30	174	9.0	622	13	AAV27532	Plasmodium falcipa
31	174	9.0	622	16	AAW68840	Plasmodium falcipa
32	145.5	7.6	354	21	AAW10640	Human VEGF-X prote
33	145.5	7.6	354	21	AAW10641	Human VEGF-X prote
34	144	7.5	49	21	AAW36293	Human GIL-19/AE289
35	134	7.0	402	8	AAV70709	Plasmodium cynomol
36	131	6.8	22	21	AAV70295	Plasmodium falcipa
37	131	6.8	54	14	AAW41356	MSP1EGF2A EGF2-lik
38	130.5	6.8	563	12	AAV10551	Plasmodium knowles
39	130	6.8	390	21	AAV49255	N19 polypeptide ca
40	128.5	6.7	66	16	AAW66442	Plasmodium falcipa
41	128	6.7	59	22	AAW63513	A peptide which ma
42	128	6.7	63	12	AAW14263	Immunogenic branch
43	127	6.6	23	21	AAV70282	Plasmodium falcipa
44	127	6.6	559	11	AAW05427	Circumsporozoite (
45	126.5	6.6	40	16	AAW87213	P.falciparum deriv

## ALIGNMENTS

RESULT 1

AAV70278  
ID AAV70278 standard; Protein: 350 AA.

XX AC AAV70278;

DT 06-JUN-2000 (first entry)

XX XX Recombinant vaccine CDC/NIIMALVAC-1.

XX KW Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;  
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;  
KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;  
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1;  
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;  
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;  
KW pf927; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody;  
KW honey bee.

XX OS Chimeric - Apis sp.  
OS Chimeric - Clostridium tetani.  
OS Chimeric - Plasmodium falciparum.

XX XX Key Peptide Location/Qualifiers

FT FT 1..22 /label= Melittin\_signal\_peptide  
FT FT /note= Derived from Honey bee  
FT FT 23..350 /label= Mature\_CDC/NIIMALVAC-1

XX XX /note= Recombinant multivalent malarial vaccine"

PN W0200011179-A1.

XX 02-MAR-2000.

PD

XX 19-AUG-1999; 99WO-US18869.  
 XX 21-AUG-1998; 98US-0097703.  
 XX (NAIM-) NAT INST IMMUNOLOGY.  
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 XX Lal AA, Shi YP, Hasnain SE;  
 XX WPI: 2000-237654/20.  
 XX N-PSDB; AA251336.  
 XX Novel recombinant protein as vaccine for treating malarial infection  
 XX comprises antigenic peptides obtained from different stages of  
 XX plasmodium falciparum life cycle  
 XX Claim 3: Page 43-44; 52pp; English.  
 XX The present sequence is that of recombinant protein CDC/N1MALVAC-1,  
 XX which is a multivalent, multi-stage malarial vaccine. The recombinant  
 XX protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope  
 XX from tetanus toxoid and 21 antigenic epitopes from circumsporozoite  
 XX protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage  
 XX antigen-1 (USA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical  
 XX membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),  
 XX rophtry associated protein-1 (RAP-1) and gamete specific antigen, Pf927.  
 XX These epitopes were obtained at different stages of the life cycle of  
 XX Plasmodium falciparum. CDC/N1MALVAC-1 vaccine has antiparasitic  
 XX activity and can be used for treatment and prevention of malarial  
 XX infections. Anti-CDC/N1MALVAC-1 antibodies can be used for detecting  
 XX P. falciparum in biological samples.  
 XX Sequence 350 AA;

Query Match 100.0%; Score 1923; DB 21; Length 350;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-151; Indels 0; Gaps 0;  
 Matches 350; Conservative 0; Mismatches 0;  
 QY 1 MKFLVNVALVFMVYISYIYADHHHKKKLUKQPGDGNPWSVCSVFCGPKDELDYEN 60  
 DB 1 mkflvnvalvfmvviyiyadhhhhkhkklkpgdgnpwspsvctcgpkdeidyn 60  
 QY 61 DIEKKICKMEKSSVFNVNSGCFRHLDERECKLLEDGSGNGKKTCECTKPDSPK 120  
 DB 61 diekkickmekssvfnvnsngcfrhlderecklledsgngkktcectkpdskp 120  
 QY 121 IVQYDNEFNPANPNANPDGNCEDIPHVNEFSAIDLGNAEYDKMDPQHYGKSLTPLE 180  
 DB 121 ivqydnfnanpnanpndgncediphvnefsaidlgnaeaydkmdpdygksltple 180  
 QY 181 ELYKPNKSLYQYIKANSKFTICITELNTFTNAGOGHMHGNEREDERTLTKEYEDIVL 240  
 DB 181 elykpndkslyqyikanskftitelntftinnagghmhghneredertltkeyedivl 240  
 QY 241 KEFTYIMNFCRGONYWEHPYQKSDQPKYQGHULTDYEEKTKGKPLDRGNYDYHYEHS 300  
 DB 241 keftymnfgrgonywehpyqksdqpkqyqghultdyektkgkpldrknydyhyehss 300  
 QY 301 PSSTKSSPNVKSASATRLMKFKAPIRDFGIGSYIEKVLAKYKDDLE 350  
 DB 301 psstkspsnvksasatrlmkfkapirodfgigisyiekvlyakykddle 350

RESULT 2

AAAR13175

IC AAAR13175 standard; Protein; 309 AA.

XX AAAR13175;

XX AAAR13175;

XX 29-AUG-1991 (first entry)

XX

DE NSL\_81-RLfdelta9.  
 XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;  
 XX hybrid; influenza virus; non-structural protein 1; fusion.  
 XX Plasmodium falciparum.  
 OS Influenza virus (A/PR/8/34/).  
 XX Key Location/Qualifiers  
 XX Region 1..81  
 FT /label= N-terminal of NSI  
 FT /note= "Influenza virus nonstructural protein 1"  
 FT Peptide 82..87  
 FT /label= synthetic linker  
 FT Region 88  
 FT /label= artifact  
 FT /note= "see comments"  
 FT Region 89..193  
 FT /label= AAs 19-123 of CS protein  
 FT /note= "Region 1 contg. flanking region less  
 FT signal sequence"  
 FT Region 194..309  
 FT /label= AAs 297-412 of CS protein  
 FT /note= "Region 11 flanking region minus 9 N-term-  
 FT inal AAs"  
 XX EP432965-A.  
 XX 19-JUN-1991.  
 XX 06-DEC-1990; 90EP-0313257.  
 XX 08-DEC-1989; 89US-0447746.  
 XX (SMIK) SMITHKLINE BEECHAM.  
 XX (USSA) US SEC OF THE ARMY.  
 XX (BIOM-) BIOMEDICAL RES INST.  
 XX Gross MS, Gordon DM, Hollingdale MR;  
 XX WPI: 1991-179771/25.  
 XX Polypeptide comprising immunogenic determinants from P falciparum  
 XX - for vaccine against malaria infection in humans.  
 XX Example 1; Page 7; 18pp; English.  
 XX The polypeptide is prepd. by genetic engineering of genes encoding  
 XX the P. falciparum circumsporozoite (CS) protein [Dame et al.,  
 XX Science 225 : 593 (1984)], and the influenza virus non-structural  
 XX protein 1 (NSI), [Baez et al., Nucleic Acids Research, 8 : 5845  
 XX (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NSI  
 XX (NSI-81) is linked via a synthetic sequence to DNA encoding Region  
 XX 1 contg. flanking regionless the 18 AA signal region, which in  
 XX turn is fused to DNA encoding Region 11-contg. flanking region  
 XX less the first nine N-terminal AAs. This CS fusion is designated  
 XX RLfdelta9. The Pro residue separating the Asp (at the C-terminal  
 XX of the linker) from RLfdelta9 is an artifact of a filled in BamHI  
 XX site. The peptide can be used in a vaccine for protection against  
 XX malaria.  
 XX See also AAR12306-RL2311 and AAR13176-RL3179.  
 XX Sequence 309 AA;

Query Match 13.1%; Score 252; DB 12; Length 309;

Best Local Similarity 45.0%; Pred. No. 7.8e-13;

Matches 54; Conservative 1; Mismatches 1; Indels 64; Gaps 2;

QY 29 KHKKLKQPGDGNP----- 41

| | | | | | | | | |

DB 179 khkklkqpgdgnpdpghmpdnrvdenanannvknneepsdkhiegylkkikns 238

Qy 42 -----WSPCSVTCG-----RPKDELVDYNDIEKKICKMEKCSSVFNVNNSG 84  
Db 239 lstwspcsvtcngigvrikpgsankpkdelyndiekkickmekcssvfnvnnsig 298

RESULT 3  
AAR07945  
XX AAR07945 standard; protein: 319 AA.  
AC AAR07945;  
XX

DT 22-FEB-1991 (first entry)  
XX NS181RLFAuth plasmid product.  
DE Malaria; vaccine.  
KW Plasmodium falciparum.  
XX

OS Plasmodium falciparum.  
XX  
FH Key Location/Qualifiers  
FT Domain 1..81  
FT /label= NS181 protein fragment  
FT /note= "from plasmid pMG-1"

FT Domain 89..193  
FT /label= Fragment of circumsporozoite protein  
FT Domain 204..319  
FT /label= Fragment of circumsporozoite protein  
XX EP398540-A.  
XX

PD 22-NOV-1990.  
XX  
XX 01-MAY-1990; 90EP-0304720.  
XX  
XX 03-MAY-1989; 89US-0346863.  
XX

XX (SMIK ) SMITHKLINE BEECHAM.  
XX  
XX Gross MS, Young JF;  
XX  
XX WPI: 1990-350299/47.  
XX N-PSDB; AAQ06580.

XX New polypeptide used in malaria vaccine - comprises immunogenic  
PT determinants from 1st and 2nd flanking regions of plasmodium  
PT surface protein and intermediate repeat domain  
XX  
XX Example 2; Page 11-12; 24pp; English.

XX The product is useful in preparation of vaccines for treatment and  
CC prophylaxis of plasmodium sporozite infection. It may be easily  
CC produced in large pure quantities from a transformed E.coli  
CC expression system.  
XX  
XX Sequence 319 AA;

PS Query Match 12.88; Score 247; DB 11; Length 319;  
CC Best Local Similarity 41.59; Pred. No. 2; le-12;  
CC Matches 54; Conservative 1; Mismatches 1; Indels 74; Gaps 2;

Qy 29 KKKLKPQGDGNP-----  
Db 179 khkkkpgdgnpdpnknngnggmmppndprnvdenanannavknnneepsdkhi 238

Qy 42 -----WSPCSVTCG-----  
Db 239 eqy1kkknstswspcsvtcngigvrikpgsankpkdelyndiekkickmekcss 298

Qy 75 VFNVNNSG 84  
Db 299 vfnvnnsig 308

Qy 42 -----WSPCSVTCG-----RPKDELVDYNDIEKKICKMEKCSS 74  
Db 239 eqy1kkknstswspcsvtcngigvrikpgsankpkdelyndiekkickmekcss 298

Qy 75 VFNVNNSG 84  
Db 299 vfnvnnsig 308

1 6 CSP

RESULT 4  
AAR13176  
ID AAR13176 standard; Protein: 319 AA.  
XX  
AC AAR13176;  
XX  
DT 29-AUG-1991 (first entry)  
XX  
DE NS1\_81-RLFAuth.  
XX

KW Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;  
KW hybrid; influenza virus; non-structural protein 1; fusion.  
XX  
OS Plasmodium falciparum.  
OS Influenza virus (A/PR/8/34).

XX Key Location/Qualifiers  
FH Region 1..81  
FT /label= N-terminal of NS1  
FT /note= "Influenza virus nonstructural protein 1"

FT Peptide 82..87  
FT /label= synthetic linker  
FT Region 88  
FT /label= artifact  
FT /note= "see comments"

FT Region 89..193  
FT /label= AAS 19-123 of CS protein  
FT /note= "Region I contg. flanking region less  
FT signal sequence"

FT Region 194  
FT /label= artifact  
FT /note= "see comments"

FT Region 195..319  
FT /label= AAS 288-412 of CS protein  
FT /note= "Region II flanking region"

XX EP432965-A.  
XX  
XX 19-JUN-1991.  
XX  
XX 06-DEC-1990; 90EP-0313257.  
XX

XX 08-DEC-1989; 89US-0447746.  
XX (SMIK ) SMITHKLINE BEECHAM.  
XX (USSA ) US SEC OF THE ARMY.  
XX (BIOM-) BIOMEDICAL RES INST.

XX Gross MS, Gordon DM, Hollingdale MR;  
XX  
XX WPI: 1991-179771/25.  
XX

XX Polypeptide comprising immunogenic determinants from P falciparum  
XX - for vaccine against malaria infection in humans.  
XX  
XX Example 2; Page 10; 18pp; English.

XX The polypeptide is prepd. by genetic engineering of genes encoding  
CC the P. falciparum circumsporozoite (CS) protein [Dame et al.,  
CC Science 225 : 593 (1984)], and the influenza virus non-structural  
CC protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845  
CC (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1  
CC (NS1-81) is linked via a synthetic sequence to DNA encoding Region  
CC I contg. flanking regionless the 18 AA signal region, which in  
CC turn is fused to DNA encoding Region II-contg. flanking region.  
CC This CS fusion is designated RLFAuth. The pro residue separating  
CC the Asp (at the C-terminal of the linker) from RLFAuth is an arti-  
CC fact of a filled in BamHI site; the Gly separating Region I and  
CC Region II-contg. CS flanking regions is an artifact of a synthetic  
CC FokI/rthIII I linker. The peptide can be used in a vaccine for  
CC protection against malaria.  
CC The complete nucleotide and AA sequences are given in EP-304720,

```

CC filed May 1, 1990.
CC See also AAR12306-R12311 and AAR13175-R13179.
XX
XX Sequence 319 AA;
SQ
    Query Match      12.8%; Score 247; DB 12; Length 319;
    Best Local Similarity 41.5%; Pred. No. 2.1e-12;
    Matches 54; Conservative 1; Mismatches 1; Indels 74; Gaps 2;

QY 29 KHKLLKQPGDGNP----- 41
Db 179 khkllkpgdgnpdknngngggghmpdpnrnvdennanannavknnneepsdkhi 238
QY 42 -----WSPCSVTCG-----KPKDELDYENDIEKKICKMEKCSS 74
Db 239 eqylkkiknsistewspcsvtcgngivrikpgsankpkdeldyendiekkickmekcss 298
QY 75 VFNVVNSNG 84
Db 299 vfnvnssig 308

RESULT 5
AAR13178
ID AAR13178 standard; Protein; 335 AA.
XX
XX AAR13178;
AC
XX
XX 29-AUG-1991 (first entry)
XX
XX NS1_81(NANP)4RLfAuth.
KW Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
KW hybrid; influenza virus; non-structural protein 1; fusion.
XX
XX
OS Plasmodium falciparum.
OS Influenza virus (A/PR/8/34).
XX
XX Key Location/Qualifiers
XX Region 1..81
XX /label= "N-terminal of NS1
XX /note= "Influenza virus nonstructural protein 1"
XX Region 82..97
XX /label= immunodominant repeat region
XX /note= "four tetrapeptide repeat units"
XX Peptide 98..103
XX /label= synthetic linker
XX Region 104
XX /label= artifact
XX /note= "see comments"
XX Region 105..209
XX /label= AAs 19-123 of CS protein
XX /note= "Region I contg. flanking region less
XX signal sequence"
XX Region 210
XX /label= artifact
XX /note= "see comments"
XX Region 211..335
XX /label= AAs 288-412 of CS protein
XX /note= "Region II flanking region"
XX
XX EP432965-A.
XX
XX 19-JUN-1991.
XX
XX 06-DEC-1990; 90EP-0313257.
XX
XX 08-DEC-1989; 89US-0447746.
XX (SMIK ) SMITHKLINE BEECHAM.
XX (USSA ) US SEC OF THE ARMY.
XX (BIOM*) BIOMEDICAL RES INST.
PA

```

---

```

Gross MS, Gordon DM, Hollingdale MR;
WPI; 1991-179771/25.
Polyptide comprising immunogenic determinants from P falciparum
- for vaccine against malaria infection in humans.
Example 4; Page 11; 18pp; English.
The polypeptide is prepd. by genetic engineering of genes encoding
the P. falciparum circumsporozoite (CS) protein (Dame et al.,
Science 225 : 593 (1984)), and the influenza virus non-structural
protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
(1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
(NS1_81) is linked to a synthetic sequence encoding four repeat
units from the immunodominant region, which in turn is linked via
a synthetic sequence to DNA encoding Region I contg. flanking
region less the 18 AA signal region. This is linked to DNA
encoding Region II contg. flanking region. The pro residue sep-
arating the Asp (at the C-terminal of the linker) from the Region
I-contg. CS flanking region is an artifact of a filled-in BamHI
site; the Gly separating the Region I and II-contg. CS flanking
regions is an artifact of a synthetic FokI/NotI linker. The
peptide can be used in a vaccine for protection against malaria.
See also AAR12306-R12311 and AAR13175-R13179.
XX
XX Sequence 335 AA;
SQ
    Query Match      12.8%; Score 247; DB 12; Length 335;
    Best Local Similarity 41.5%; Pred. No. 2.3e-12;
    Matches 54; Conservative 1; Mismatches 1; Indels 74; Gaps 2;

QY 29 KHKLLKQPGDGNP----- 41
Db 195 khkllkpgdgnpdknngngggghmpdpnrnvdennanannavknnneepsdkhi 254
QY 42 -----WSPCSVTCG-----KPKDELDYENDIEKKICKMEKCSS 74
Db 255 eqylkkiknsistewspcsvtcgngivrikpgsankpkdeldyendiekkickmekcss 314
QY 75 VFNVVNSNG 84
Db 315 vfnvnssig 324

RESULT 6
AAR13179
ID AAR13179 standard; Protein; 335 AA.
XX
XX AAR13179;
XX
XX 29-AUG-1991 (first entry)
XX
XX NS1_81(NVDP)4RLfAuth.
KW Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
KW hybrid; influenza virus; non-structural protein 1; fusion.
XX
XX Plasmodium falciparum.
XX Influenza virus (A/PR/8/34).
XX
XX Key Location/Qualifiers
XX Region 1..81
XX /label= "N-terminal of NS1
XX /note= "Influenza virus nonstructural protein 1"
XX Region 82..97
XX /label= immunodominant repeat region
XX /note= "four variant tetrapeptide repeat units"
XX Peptide 98..103
XX /label= synthetic linker
XX Region 104

```



Mon Feb 4 15:23:41 2002

CC peptide can be used in a vaccine for protection against malaria.  
 CC See also AAR12306-R12311 and AAR13175-R13179.  
 XX  
 SQ Sequence 327 AA;

Query Match 12.6%; Score 243; DB 12; Length 327;  
 Best Local Similarity 39.1%; Pred. No. 4.7e-12;  
 Matches 54; Conservative 1; Mismatches 1; Indels 82; Gaps 2;  
 QY 29 KHKLLKQPGGNGP-----  
 Db 179 khkklkpggngpdpnanpanpnknkngngggghmpdpnrnvdenanannavknmm 238  
 QY 42 -----WSPCSVTG-----KPKDELDYENDIEKKI 66  
 Db 239 eepdkhiegylkkiknsistewspcsvtcgnglqvrikpgsankpkdelidyendiekk 298  
 QY 67 CKMEKCSVFNVSNSG 84  
 Db 299 ckmekcsvfnvnssig 316

RESULT 8

AAB37609  
 ID AAB37609 standard; protein; 108 AA.

XX  
 AC AAB37609;

XX 27-FEB-2001 (first entry)

XX Merozoite surface protein-119.

XX Merozoite surface protein; protazoacide; vaccine; malaria.

XX Plasmodium falciparum.

XX WO200063245-A2.

XX 26-OCT-2000.

XX 20-APR-2000; 2000WO-GB01558.

XX 20-APR-1999; 99GB-0009072.

XX 13-MAY-1999; 99US-0311817.

XX 25-MAY-1999; 99CA-2271451.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Holder A, Birdsall B, Feeney J, Morgan W, Syed S, Uthaiipbull C;

XX WPI: 2001-015762/02.

XX N-PSDB; AAC68977.

XX Novel variants of the C-terminal fragment of Plasmodium merozoite

XX surface protein-1, useful as vaccines for treating or preventing

XX malaria

XX Example 5; Fig 15; 126pp; English.

XX The present invention relates to non-natural variants of a C-terminal  
 CC fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The  
 CC non-natural variants have reduced affinity for at least 1 antibody  
 CC capable of blocking a second antibody that inhibits the proteolytic  
 CC cleavage of Plasmodium MSP-1.4.2, and has the same affinity for at least  
 CC one third antibody that inhibits the proteolytic cleavage of Plasmodium  
 CC MSP-1.4.2, compared to natural MSP-1.1.9. The non-natural variants of the  
 CC present invention are useful for immunising a mammal against malaria, and  
 CC can be used to treat malaria. The present sequence is MSP-119 protein.

XX Sequence 108 AA;

Query Match 11.5%; Score 222; DB 22; Length 108;  
 Best Local Similarity 35.9%; Pred. No. 5.4e-11;  
 Matches 47; Conservative 8; Mismatches 10; Indels 66; Gaps 3;

QY 23 .HHHHHHKHLKQPCDGNPWSVTCGPKDELIDYENDIEKKICKMEKCSVFNVSNS 82  
 Db 2 hhhhhh-----legrhniaqhgcvkqcp-----qp 27  
 QY 83 SGCFRHLDERECKLL-----EDSGSNGKKIT 110  
 Db 28 sgcfhrldereckllnykqgdkcvenpntcnennngcdadakteedsgsngkkit 87  
 QY 111 CECTKPDSPKI 121  
 Db 88 cectkpdspyl 98

RESULT 9

AAP83144  
 ID AAP83144 standard; protein; 411 AA.

XX AAP83144;

XX 20-NOV-1990 (first entry)

XX Sequence encoded by the circumsporozoite (CS) gene from

XX Plasmodium falciparum.

XX Vaccine; antigen; immunogen; probe; hybridisation;

XX immunoassay; diagnosis.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

XX Region 106..120

XX /note="Region 1"

XX Region 147..206

XX /note="Repeat region, repeat unit = NANP"

XX Region 123..146

XX /note="Repeat region, repeat unit = NANPNDP"

XX Region 211..286

XX /note="Repeat region, repeat unit = NANP"

XX EP278940-A.

XX 17-AUG-1988.

XX 25-JAN-1988; 88EP-0870008.

XX 30-JAN-1987; 87US-0009325.

XX (SMIK ) SMITH KLINE-RIT.

XX Cabazon T, De Wilde M, Harford N;

XX WPI: 1988-229751/33.

XX N-PSDB; AAN81108.

XX DNA encoding hepatitis B virus antigens and hybrids contg. them -

XX used for expression in yeast to obtain vaccines and bivalent

XX vaccines

XX Example; Fig 3Aa-3Af; 101pp; English.

XX Sequence of the CS gene (AAN81108) is from lambda-mbfl. A recombinant

XX DNA molecule is claimed, comprising functional DNA coding sequence fused,

XX in phase, to a portion of the Pre-S2 region of a hepatitis B virus (HBV)

XX pre-S2-S protein coding sequence. The functional DNA coding sequence

XX comprises the Pre-S2 coding sequence, the Pre-S1 coding sequence or entire

XX pre-S1-Pre-S2 coding sequence, the CS protein coding sequence of

XX Plasmodium, or a HIV coding sequence such as

XX an HIV envelope gene sequence, e.g. HIV C7 protein coding region, HIV

XX Peptide 121 coding region, or HIV Dreesman peptide coding region.



XX  
SQ Sequence 411 AA;

Query Match 11.28; Score 215.5; DB 9; Length 411;

Best Local Similarity 62.28; Pred. No. 1.2e-09;

Sequence	Length	Matches	Mismatches	Indels	Gaps
...	...	46	2	7	19
...	...	46	2	7	19

QY 29 KH-----KKLKQPGDGNPWSPCSVTCG-----KPKDELDYENDIEKKICKME 70

[illegible]

QY 71 KCSSVENVNSNG 84

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

**RESULT 10**

AAP60416  
 ID AAP60416 standard; Protein; 412 AA.

AC AAP60416:

DT 13-JUN-1991 (first entry)

CS protein of malarial parasite.

XX Sporozoite: vaccination.

XX Plasmodium falciparum. OS

XX	Key	Location/Qualifiers
FH	Region	124..127
FT		/label= Repeat unit
FT		

FT	Region	Location
124	127	127

FT /label= Repeat unit

PN EP166410-A.

02-JAN-1986.

XX  
PF 24-JUN-1985: 85EP-0107794.

XX 26-JUN-1984: 84US-0624564-

XX  
PA (USDC ) IIS SEC OF COMMERCE

PA (USDC ) US SEC OF COMMERCE.  
PA (USGO ) IIS GOVERNMENT

PA (USGU ) US GOVERNMENT.  
PA (USSA ) US SEC OF THE ARMY.

XX PI McCutchan TF, Dame JB, Williams JL, Schneider T:

XX  
DB  
WPT: 1986-008635/02

DR WP1: I986-00863570  
DR N-PSDB: AAN60362

xx New immunologically active pure synthetic peptide(s) - used for  
PT protection against infection by malaria parasite.  
PT

PS Disclosure; Fig 2: 49pp: English.

The Plasmodium CS gene was used to isolate peptides capable of inducing an immune response to the parasite. Peptide antigens may be synthesised in pure form and used to generate an immune response in vaccination against malaria. The featured repeat units are claimed and must be present in copies of 2-1000.

XX	Sequence	412 AA:
SO		

### Query Match

Query Match 11.2%; Score 213.3; DB 7; Length 412;  
Best Local Similarity 62.2%; Pred. No. 1.2e-09;

Matches	46; Conservative	2; Mismatches	7
---------	------------------	---------------	---

QY 29 KH-----KKLKOPGDGNPWSPCSVTCG-----KPKDELDYENDIEKKICKME 70

Q	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100																								
Q	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

Query Match 11.0%; Score 211.5; DB 9; Length 412;  
 Best Local Similarity 60.8%; Pred. No. 2.6e-09;  
 Matches 45; Conservative 3; Mismatches 7; Indels 19; Gaps 3;

QY 29 KH-----KKLKQPGDGNPSPCSVTCG-----KPKDELVDYENDIEKKICKME 70  
 || ||:| |||||  
 Db 329 khlegylkkikn-sistewspcsvtcngiqvrikpsankpkdelyndiekkickie 387

QY 71 KCSSVFNVNSG 84  
 |||||  
 Db 388 kcsvgfnvnssig 401

RESULT 12  
 AAR37796  
 ID AAR37796 standard; Protein: 424 AA.  
 XX  
 AC AAR37796;  
 DT 27-SEP-1993 (first entry)  
 XX  
 DE RTS protein.  
 XX  
 KW RTS; expression cassette; hybrid protein; S. cerevisiae; TDH3;  
 KW cloning; circumsporozoite protein; CSP; Plasmodium falciparum;  
 KW strain 7G8; hepatitis B virus; HBV; adw serotype; pres2 protein;  
 KW S protein.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1  
 FT /note= "Derived from S. cerevisiae TDH3 gene sequence"  
 FT Region 2..4  
 FT /note= "Cloning artefact"  
 FT Protein 5..193  
 FT /note= "Represents amino acids 210-398 of the CSP of  
 FT P. falciparum"  
 FT Region 194..197  
 FT /note= "Carboxy terminal amino acids from HBV (adw  
 FT serotype) pres2 protein"  
 FT Protein 198..424  
 FT /note= "S protein of HBV (adw serotype)"  
 FT  
 FT WO9310152-A.  
 PN  
 XX  
 PD 27-MAY-1993.  
 XX  
 PF 11-NOV-1992; 92WO-EP02591.  
 XX  
 PR 16-NOV-1991; 91GB-0024390.  
 PR 27-FEB-1992; 92US-0842694.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Cohen J, De Wilde M;  
 XX  
 DR WPI: 1993-182494/22.  
 DR N-PSDB; AAQ42566.  
 XX  
 XX Hybrid protein comprising Plasmodium circumsporozoite protein and  
 PT HBSAg - useful as a vaccine for treating patients susceptible to  
 PT Plasmodium infections  
 PT  
 XX Disclosure; Fig 5; 59pp; English.  
 PS  
 XX This sequence represents the RTS hybrid protein which is encoded by  
 CC the RTS expression cassette. This hybrid consists of a methionine  
 CC residue derived from S. cerevisiae TDH3 gene sequence, three amino  
 CC acids, Met-Ala-Pro, derived from a nucleotide sequence created by  
 CC the cloning procedure used to construct the hybrid gene, a stretch  
 CC of 189 amino acids representing amino acids 210 to 398 of the

CC circumsporozoite protien (CSP) of Plasmodium falciparum strain 7G8,  
 CC an amino acid Arg created by the cloning procedure, four amino acids,  
 CC pro-Val-Thr-Asn, representing the four carboxy terminal residues of  
 CC hepatitis B virus (HBV), adw serotype, pres2 protein, and a stretch  
 CC of 226 amino acids specifying the S protein of HBV, adw serotype.  
 CC This protein, and RTS\* (see also AAR37797), may be combined with an  
 CC adjuvant and used in a vaccine for preventing Plasmodium infections.  
 CC The vaccines produce a humoral response and also a cellular immune  
 CC response.  
 XX  
 SQ Sequence 424 AA;  
 Query Match 10.9%; Score 210.5; DB 14; Length 424;  
 Best Local Similarity 63.4%; Pred. No. 3.3e-09;  
 Matches 45; Conservative 1; Mismatches 6; Indels 19; Gaps 3;

QY 29 KH-----KKLKQPGDGNPSPCSVTCG-----KPKDELVDYENDIEKKICKME 70  
 || ||:| |||||  
 Db 124 khlegylkkikn-sistewspcsvtcngiqvrikpsankpkdelyndiekkickme 182

QY 71 KCSSVFNVNSG 81  
 |||||  
 Db 183 kcsvgfnvnns 193

RESULT 13  
 AAR37797  
 ID AAR37797 standard; Protein: 424 AA.  
 XX  
 AC AAR37797;  
 DT 27-SEP-1993 (first entry)  
 XX  
 DE RTS\* protein.  
 XX  
 KW RTS; expression cassette; hybrid protein; S. cerevisiae; TDH3;  
 KW cloning; circumsporozoite protein; CSP; Plasmodium falciparum;  
 KW strain 7G8; hepatitis B virus; HBV; adw serotype; pres2 protein;  
 KW S protein.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1  
 FT /note= "Derived from S. cerevisiae TDH3 gene sequence"  
 FT Region 2..4  
 FT /note= "Cloning artefact"  
 FT Protein 5..193  
 FT /note= "Represents amino acids 210-398 of the CSP of  
 FT P. falciparum"  
 FT Region 194..197  
 FT /note= "Carboxy terminal amino acids from HBV (adw  
 FT serotype) pres2 protein"  
 FT Protein 198..424  
 FT /note= "S protein of HBV (adw serotype)"  
 FT  
 FT WO9310152-A.  
 PN  
 XX  
 PD 27-MAY-1993.  
 XX  
 PF 11-NOV-1992; 92WO-EP02591.  
 XX  
 PR 16-NOV-1991; 91GB-0024390.  
 PR 27-FEB-1992; 92US-0842694.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Cohen J, De Wilde M;  
 XX  
 DR WPI: 1993-182494/22.  
 DR N-PSDB; AAQ42567.  
 XX

PT Hybrid protein comprising Plasmodium circumsporozoite protein and  
PT HSAg - useful as a vaccine for treating patients susceptible to  
XX Plasmodium infections  
PS Disclosure; Fig 9; 59pp; English.  
XX This sequence represents the RTS\* hybrid protein which is encoded by  
CC the RTS\* expression cassette. This hybrid consists of a methionine  
CC residue derived from S. cerevisiae TSH3 gene sequence, three amino  
CC acids, Met-Ala-Pro, derived from a nucleotide sequence created by  
CC the cloning procedure used to construct the hybrid gene, a stretch  
CC of 189 amino acids representing amino acids 210 to 398 of the  
CC circumsporozoite protein (CSP) of Plasmodium falciparum strain NF54,  
CC an amino acid Arg created by the cloning procedure, four amino acids,  
CC pro-Val-Thr-Asn, representing the four carboxy terminal residues of  
CC hepatitis B virus (HBV), adw serotype, pres2 protein, and a stretch  
CC of 226 amino acids specifying the S1 protein of HBV, adw serotype.  
CC This protein, and RTS (see also AAR37796), may be combined with an  
CC adjuvant and used in a vaccine for preventing plasmodium infections.  
CC The vaccines produce a humoral response and also a cellular immune  
XX response.

SQ Sequence 424 AA;

Query Match 10.8%; Score 207.5; DB 14; Length 424;

Best Local Similarity 51.8%; Pred. No. 5.9e-03;

Matches 43; Conservative 4; Mismatches 9; Indels 27; Gaps 2;

QY 29 KHKLKQPGD-----GNPWSPCSVTCG-----KPKDELQYEND 61

Db 110 knnnneepsdkhikeylnkignslstewspcsvtcgnglqvrikpgsankpkdelidyaand 169

QY 62 TEKKICKMEKCSVFNVSNSG 84

Db 170 lekkickmekcsvgfnvnssig 192

RESULT 14

AAW36103

ID AAW36103 standard; Protein; 116 AA.

XX XX

AC AAW36103;

XX XX

DT 25-MAR-1998 (first entry)

XX XX

DE PfMSP1(p19)A protein sequence.

XX XX

KW Plasmodium vivax; merozoite surface protein; MSP1; p19;

XX XX

KW Plasmodium falciparum; malaria; vaccine; immunity; epitope.

XX XX

OS Plasmodium falciparum.

XX XX

OS Synthetic.

XX XX

FH Key Location/Qualifiers

FT Region 1..95

FT /note= "amino acids derived from P. falciparum MSP1 p19

FT fragment"

FT Region 96..116

FT /note= "glycosylphosphatidylinositol anchoring sequence"

XX XX

PN WO9730158-A2.

XX XX

PD 21-AUG-1997.

XX XX

PF 14-FEB-1997; 97WO-FR00290.

XX XX

PR 14-FEB-1996; 96FR-0001822.

XX XX

PA (INSP ) INST PASTEUR.

XX XX

PA (UUNY ) UNIV NEW YORK STATE.

XX XX

PI Barnwell JW, Longacre-Andre S, Mendis K, Nato F;

PI Roth C;

XX XX

DR WPI; 1997-425033/39.

XX XX

DR N-PSDB; AAT94550.

XX XX

PT Recombinant protein containing the merozoite surface protein-1 p19

PT fragment - useful in anti-malarial vaccines, diagnosis and protein

XX XX

PS Disclosure; Fig 1B; 85pp; French.

XX XX

CC This is the amino acid sequence of a recombinant protein comprising

CC amino acids 1613-1705 of the Plasmodium falciparum merozoite surface

CC protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a

CC glycosylphosphatidylinositol membrane anchoring sequence. p19 is the

CC C-terminal fragment of the 42 kD MSP1 from plasmodium species.

CC The recombinant protein can be used for the production of anti-malarial

CC vaccines, where the p19 fragment provides a high level of protective

CC immunity since it includes epitopes not presented in the p42 fragment.

XX XX

SQ Sequence 116 AA;

Query Match 10.3%; Score 198.5; DB 18; Length 116;

Best Local Similarity 44.2%; Pred. No. 5.3e-09;

Matches 42; Conservative 5; Mismatches 9; Indels 39; Gaps 2;

QY 59 ENDIEKKICKMEKCSVFNVSNSGCFRHLDERECKCLL----- 99

Db 1 efnlsgqgvkqgc-----ensgcfhldereeckcllnykggdkovenpnptcne 53

QY 100 -----EDSGNSGKKTCTCTKPDSPKPI 121

Db 54 nnggcdadaktcedsgngkkitcectkpdspyl 88

RESULT 15

AAW22592

ID AAW22592 standard; Protein; 116 AA.

XX XX

AC AAW22592;

XX XX

DT 25-MAR-1998 (first entry)

XX XX

DE PfMSP1(p19)A protein sequence.

XX XX

KW Plasmodium vivax; merozoite surface protein; MSP1; p19;

XX XX

KW Plasmodium falciparum; malaria; vaccine; immunity; epitope.

XX XX

OS Plasmodium falciparum.

XX XX

OS Synthetic.

XX XX

FH Key Location/Qualifiers

FT Region 1..95

FT /note= "amino acids derived from P. falciparum MSP1 p19

FT fragment"

FT Region 96..116

FT /note= "glycosylphosphatidylinositol anchoring sequence"

XX XX

PN WO9730159-A2.

XX XX

PD 21-AUG-1997.

XX XX

PF 14-FEB-1997; 97WO-FR00291.

XX XX

PR 14-FEB-1996; 96FR-0001821.

XX XX

PA (INSP ) INST PASTEUR.

XX XX

PA (UUNY ) UNIV NEW YORK STATE.

XX XX

PI Barnwell JW, Longacre-Andre S, Mendis K, Nato F;

XX XX

PI Roth C;



Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	196.5	10.2	96	5	US-09-978-756-1	Sequence 1, Appli	
2	196.5	10.2	376	5	US-09-978-756-2	Sequence 2, Appli	
3	196.5	10.2	594	5	US-09-117-415B-2	Sequence 2, Appli	
4	196.5	10.2	613	5	US-09-117-415B-22	Sequence 22, Appl	
5	196.5	10.2	631	5	US-09-117-415B-18	Sequence 18, Appl	
6	196.5	10.2	631	5	US-09-117-415B-20	Sequence 20, Appl	
7	196.5	10.2	649	5	US-09-117-415B-16	Sequence 16, Appl	
8	196.5	10.2	1602	5	US-09-269-874A-7	Sequence 7, Appli	
9	196.5	10.2	1621	5	US-09-269-874A-5	Sequence 5, Appli	
10	196.5	10.2	1639	5	US-09-269-874A-3	Sequence 3, Appli	
11	187.5	9.8	393	5	US-09-978-756-3	Sequence 3, Appli	
12	109.5	5.7	905	5	US-09-445-096A-70	Sequence 70, Appl	
13	108	5.6	1189	5	US-09-402-181A-613	Sequence 613, App	
14	108	5.6	1189	5	US-09-721-477-613	Sequence 613, App	
15	105.5	5.5	286	5	US-09-708-427-30543	Sequence 30543, A	
16	104	5.4	484	5	US-09-708-427-13368	Sequence 13368, A	
17	104	5.4	486	5	US-09-708-427-13367	Sequence 13367, A	
18	104	5.4	516	5	US-09-708-427-14308	Sequence 14308, A	
19	104	5.4	518	5	US-09-708-427-14307	Sequence 14307, A	
20	104	5.4	707	5	US-09-708-427-13366	Sequence 13366, A	
21	104	5.4	739	5	US-09-708-427-14306	Sequence 14306, A	
22	102	5.3	173	5	US-09-787-126-18	Sequence 18, Appl	
23	98.5	5.1	716	6	US-10-001-05A-30	Sequence 30, Appl	
24	98.5	5.1	3418	5	US-09-084-471C-5	Sequence 5, Appli	
25	98.5	5.1	3418	5	US-09-084-471C-7	Sequence 7, Appli	
26	98.5	5.1	3418	5	US-09-084-471C-9	Sequence 9, Appli	

APPLICANT: Feeney, James  
APPLICANT: Morgan, William  
APPLICANT: Syed, Shabih  
TITLE OF INVENTION: Malaria Vaccine  
FILE REFERENCE: 18396/1005  
CURRENT APPLICATION NUMBER: US/09/978,756  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: PCT/GB00/01558  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: 09/311,817  
PRIOR FILING DATE: 1999-05-13  
PRIOR APPLICATION NUMBER: 2,271,451  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 9909072.2  
PRIOR FILING DATE: 1999-04-20  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 376  
TYPE: PRT  
ORGANISM: Plasmodium falciparum  
US-09-978-756-2

Query Match 10.2%; Score 196.5; DB 5; Length 376;  
Best Local Similarity 44.1%; Pred. No. 8.7e-08;  
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;  
QY 61 DIEKKICKMEKCSSVFNVNSGCFRHLDERECKCLL-----99  
Db 263 NISQHCVKKQCP-----QNSGCFRHLDERECKCLLYKQEGDKCVENPNPTCNENN 315  
QY 100 -----EDSGNGKKTCTCKTPDPSKPI 121  
Db 316 GGCDADAKCTEEDSGSGNGKKTCTCKTPDPSYPL 348

RESULT 3  
US-09-117-415B-2  
Sequence 22 Application US/09117415B  
GENERAL INFORMATION:  
APPLICANT: Davidson, Eugene  
Yang, Shutong  
TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
of a MSAL Peptide  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
STREET: 714 Colorado Avenue  
CITY: Bridgeport  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 10017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPad (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/117,415B  
FILING DATE: 29-Jul-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coleman, Henry D.  
REGISTRATION NUMBER: 32,559  
REFERENCE/DOCKET NUMBER: R12-030  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 679-0090  
TELEFAX: (212) 679-9121  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 594 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-117-415B-2  
Query Match 10.2%; Score 196.5; DB 5; Length 594;  
Best Local Similarity 44.1%; Pred. No. 1.5e-07;  
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;  
QY 61 DIEKKICKMEKCSSVFNVNSGCFRHLDERECKCLL-----99  
Db 481 NISQHCVKKQCP-----QNSGCFRHLDERECKCLLYKQEGDKCVENPNPTCNENN 533  
QY 100 -----EDSGNGKKTCTCKTPDPSKPI 121  
Db 534 GGCDADAKCTEEDSGSGNGKKTCTCKTPDPSYPL 566

RESULT 4  
US-09-117-415B-22  
Sequence 22 Application US/09117415B  
GENERAL INFORMATION:  
APPLICANT: Davidson, Eugene  
Yang, Shutong  
TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
of a MSAL Peptide  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
STREET: 714 Colorado Avenue  
CITY: Bridgeport  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 10017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPad (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/117,415B  
FILING DATE: 29-Jul-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coleman, Henry D.  
REGISTRATION NUMBER: 32,559  
REFERENCE/DOCKET NUMBER: R12-030  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 679-0090  
TELEFAX: (212) 679-9121  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 613 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-117-415B-22

Query Match 10.2%; Score 196.5; DB 5; Length 613;  
Best Local Similarity 44.1%; Pred. No. 1.6e-07;  
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;  
QY 61 DIEKKICKMEKCSSVFNVNSGCFRHLDERECKCLL-----99  
Db 518 NISQHCVKKQCP-----QNSGCFRHLDERECKCLLYKQEGDKCVENPNPTCNENN 570  
QY 100 -----EDSGNGKKTCTCKTPDPSKPI 121  
Db 571 GGCDADAKCTEEDSGSGNGKKTCTCKTPDPSYPL 603

RESULT 5  
US-09-117-415B-18  
; Sequence 18, Application US/09117415B  
; GENERAL INFORMATION:  
; APPLICANT: Davidson, Eugene  
; Yang, Shutong  
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
; of a MSAL Peptide  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
; STREET: 714 Colorado Avenue  
; CITY: Bridgeport  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 10017  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPad (ASCII)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/117,415B  
; FILING DATE: 29-Jul-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coleman, Henry D.  
; REGISTRATION NUMBER: 32,559  
; REFERENCE/DOCKET NUMBER: R12-030  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 679-0090  
; TELEFAX: (212) 679-9121  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 631 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-117-415B-18  
Query Match 10.2%; Score 196.5; DB 5; Length 631;  
Best Local Similarity 44.1%; Pred. No. 1.6e-07;  
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;  
QY 61 DIEKKICKMEKCSVFNVNSGCFRHLDERECKCLL-----99  
: : : : :  
Db 536 NISQHCVKKQCP-----QNSGCFRHLDERECKCLLNRYKQEGDKCVENPNTCNENN 588  
QY 100 -----EDSGSNGKKITCECTKPDSPKI 121  
: : : : :  
Db 589 GGCDAKACTEEDSGSNGKKITCECTKPDSPYL 621  
RESULT 6  
US-09-117-415B-20  
; Sequence 20, Application US/09117415B  
; GENERAL INFORMATION:  
; APPLICANT: Davidson, Eugene  
; Yang, Shutong  
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
; of a MSAL Peptide  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
; STREET: 714 Colorado Avenue  
; CITY: Bridgeport  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 10017  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPad (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/117,415B  
FILING DATE: 29-Jul-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coleman, Henry D.  
REGISTRATION NUMBER: 32,559  
REFERENCE/DOCKET NUMBER: R12-030  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 679-0090  
TELEFAX: (212) 679-9121  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 631 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-117-415B-20  
Query Match 10.2%; Score 196.5; DB 5; Length 631;  
Best Local Similarity 44.1%; Pred. No. 1.6e-07;  
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;  
QY 61 DIEKKICKMEKCSVFNVNSGCFRHLDERECKCLL-----99  
: : : : :  
Db 518 NISQHCVKKQCP-----QNSGCFRHLDERECKCLLNRYKQEGDKCVENPNTCNENN 570  
QY 100 -----EDSGSNGKKITCECTKPDSPKI 121  
: : : : :  
Db 571 GGCDAKACTEEDSGSNGKKITCECTKPDSPYL 603  
RESULT 7  
US-09-117-415B-16  
; Sequence 16, Application US/09117415B  
; GENERAL INFORMATION:  
; APPLICANT: Davidson, Eugene  
; Yang, Shutong  
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
; of a MSAL Peptide  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
; STREET: 714 Colorado Avenue  
; CITY: Bridgeport  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 10017  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPad (ASCII)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/117,415B  
; FILING DATE: 29-Jul-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coleman, Henry D.  
; REGISTRATION NUMBER: 32,559  
; REFERENCE/DOCKET NUMBER: R12-030  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 679-0090  
; TELEFAX: (212) 679-9121  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 649 amino acids

PCT/US97/01395

LENGTH: 1621  
TYPE: PRT  
ORGANISM: Plasmodium falciparum  
US-09-269-874A-5

Query Match 10.2%; Score 196.5; DB 5; Length 1621;  
Best Local Similarity 44.1%; Pred. No. 5.1e-07;  
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;

QY 61 DIEKKICKMEKCSVFNVNSGCFRHLDERECKCLL-----99  
Db 1526 NISQHCVKKQCP-----QNSGCFRHLDERECKCLLNYKQEGDKCVENPNTCNENN 1578

QY 100 -----EDSGNGKKITCECTKPDSPKI 121  
Db 1579 GGCDAKACTEEDSGNGKKITCECTKPDSPYL 1611

RESULT 10  
US-09-269-874A-3  
Sequence 3, Application US/09269874A  
GENERAL INFORMATION:  
APPLICANT: Bujard, Hermann  
TITLE OF INVENTION: Recombinant Process for Preparing a  
FILE REFERENCE: GRUE-003  
CURRENT APPLICATION NUMBER: US/09/269,874A  
CURRENT FILING DATE: 1999-08-02  
PRIOR APPLICATION NUMBER: PCT/EP97/05441  
PRIOR FILING DATE: 1997-10-02  
PRIOR APPLICATION NUMBER: DE 19640817.2  
PRIOR FILING DATE: 1996-10-02  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 1639  
TYPE: PRT  
ORGANISM: Plasmodium falciparum  
US-09-269-874A-3

Query Match 10.2%; Score 196.5; DB 5; Length 1639;  
Best Local Similarity 44.1%; Pred. No. 5.2e-07;  
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;

QY 61 DIEKKICKMEKCSVFNVNSGCFRHLDERECKCLL-----99  
Db 1526 NISQHCVKKQCP-----QNSGCFRHLDERECKCLLNYKQEGDKCVENPNTCNENN 1578

QY 100 -----EDSGNGKKITCECTKPDSPKI 121  
Db 1579 GGCDAKACTEEDSGNGKKITCECTKPDSPYL 1611

RESULT 11  
US-09-978-756-3  
Sequence 3, Application US/09978756  
GENERAL INFORMATION:  
APPLICANT: Holder, Anthony  
APPLICANT: Birdsell, Berry  
APPLICANT: Feeney, James  
APPLICANT: Morgan, William  
APPLICANT: Syed, Shabih  
TITLE OF INVENTION: Malaria vaccine  
FILE REFERENCE: 18396/1005  
CURRENT APPLICATION NUMBER: US/09/978,756  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: PCT/GB00/01558  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: 09/311,817  
PRIOR FILING DATE: 1999-05-13  
PRIOR APPLICATION NUMBER: 2,271,451

us-09-117-415B-16

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Query Match 10.2%; Score 196.5; DB 5; Length 649;  
Best Local Similarity 44.1%; Pred. No. 1.7e-07;  
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;

QY 61 DIEKKICKMEKCSVFNVNSGCFRHLDERECKCLL-----99  
Db 536 NISQHCVKKQCP-----QNSGCFRHLDERECKCLLNYKQEGDKCVENPNTCNENN 588

QY 100 -----EDSGNGKKITCECTKPDSPKI 121  
Db 589 GGCDAKACTEEDSGNGKKITCECTKPDSPYL 621

RESULT 8  
US-09-269-874A-7  
Sequence 7, Application US/09269874A  
GENERAL INFORMATION:  
APPLICANT: Bujard, Hermann  
TITLE OF INVENTION: Recombinant Process for Preparing a  
FILE REFERENCE: GRUE-003  
CURRENT APPLICATION NUMBER: US/09/269,874A  
CURRENT FILING DATE: 1999-08-02  
PRIOR APPLICATION NUMBER: PCT/EP97/05441  
PRIOR FILING DATE: 1997-10-02  
PRIOR APPLICATION NUMBER: DE 19640817.2  
PRIOR FILING DATE: 1996-10-02  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 1602  
TYPE: PRT  
ORGANISM: Plasmodium falciparum  
US-09-269-874A-7

Query Match 10.2%; Score 196.5; DB 5; Length 1602;  
Best Local Similarity 44.1%; Pred. No. 5e-07;  
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;

QY 61 DIEKKICKMEKCSVFNVNSGCFRHLDERECKCLL-----99  
Db 1507 NISQHCVKKQCP-----QNSGCFRHLDERECKCLLNYKQEGDKCVENPNTCNENN 1559

QY 100 -----EDSGNGKKITCECTKPDSPKI 121  
Db 1560 GGCDAKACTEEDSGNGKKITCECTKPDSPYL 1592

RESULT 9  
US-09-269-874A-5  
Sequence 5, Application US/09269874A  
GENERAL INFORMATION:  
APPLICANT: Bujard, Hermann  
TITLE OF INVENTION: Recombinant Process for Preparing a  
FILE REFERENCE: GRUE-003  
CURRENT APPLICATION NUMBER: US/09/269,874A  
CURRENT FILING DATE: 1999-08-02  
PRIOR APPLICATION NUMBER: PCT/EP97/05441  
PRIOR FILING DATE: 1997-10-02  
PRIOR APPLICATION NUMBER: DE 19640817.2  
PRIOR FILING DATE: 1996-10-02  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5





```

;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; FEATURE:
;
; NAME/KEY: Protein
;
; LOCATION: 1..1189
;
; OTHER INFORMATION: /note= "fusion protein composed of
; melittin signal sequence and full length
; hTrk protein"
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 613:
US-09-721-477-613

```

```
Query Match          5.6%; Score 108; DB 5; Length 1189;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

RESULT 15
US-09-708-427-30543
; Sequence 30543, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30543
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..286
; OTHER INFORMATION: xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..286
; OTHER INFORMATION: Ceres Seq. ID 1829492
; US-09-708-427-30543

```

Query Match	5.5%;	Score 105.5;	DB 5;	Length 286;
Best Local Similarity	20.1%;	Pred. No. 0.62;		
Matches	60: Conservative	35: Mismatches	112: Indels	91: Gaps
				13:

				5.6%;	Score 108;	DB 5;	Length 1189;
Query Match							
Best Local Similarity				100.0%;	Pred. No. 2;2;		
Matches	22;	Conservative	0;	Mismatches	0;	Indels	
QY	1	MKFLVNVALFMVVIYSIYYAD	22				
Dd	1	MKELVNVALFMVVIYSIYYAD	22				

RESULT 14  
US-09-721-477-613  
; Sequence 613, Application US/09721477  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Lingner, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.  
; Morin, Gregg B.  
; Harley, Calvin B.  
; Andrews, William H.  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/09/721,477  
; APPLICATION NUMBER: US/09/721,477  
; FILING DATE: 22-Nov-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/974,549  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/911,312  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: US 08/912,951  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: US 08/915,503  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: WO PCT/US97/17618  
; FILING DATE: 01-OCT-1997  
; APPLICATION NUMBER: WO PCT/US97/17885  
; FILING DATE: 01-OCT-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph Ted  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-0026100S  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 613:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1189 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>

Mon Feb 4 15:23:41 2002

us-09-763-397a-2.rapn

Page 7

Search completed: January 29, 2002, 10:58:07  
Job time: 2503 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:15:49 ; Search time 144.96 Seconds  
(without alignments)  
183.920 Million cell updates/sec

Title: US-09-763-397A-2  
Perfect score: 1923  
Sequence: 1 MKFLVNVALFVFMVYIY.....DFFGISYKVLAKYKDDLE 350

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues  
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216.5	11.3	424	2 A54533	circumsporozoite p
2	215.5	11.2	412	1 OZZQAF	circumsporozoite p
3	213.5	11.1	442	2 A54529	circumsporozoite p
4	208.5	10.8	388	2 A39756	circumsporozoite p
5	207.5	10.8	405	2 S05428	circumsporozoite p
6	200.5	10.4	651	2 S47282	merozoite surface
7	200.5	10.4	1726	1 SAZQGM	major merozoite su
8	200.5	10.4	1726	2 A45948	major merozoite su
9	196.5	10.2	400	2 A55545	major merozoite su
10	196.5	10.2	1639	2 S05603	major merozoite su
11	196	10.2	1631	1 SAZQK1	major merozoite su
12	187.5	9.8	1701	2 A54498	major merozoite su
13	187.5	9.8	1701	2 A26868	major merozoite su
14	186	9.7	1640	2 A24594	probable major sur
15	179	9.3	622	2 A32499	apical membrane an
16	174	9.0	622	2 B44986	apical membrane an
17	172.5	9.0	622	2 A44986	apical membrane an
18	171.5	8.9	622	2 D44986	apical membrane an
19	171	8.9	622	2 C44986	apical membrane an
20	135	7.0	378	1 OZZQAB	circumsporozoite p
21	135	7.0	378	1 OZZQAL	circumsporozoite p
22	135	7.0	398	1 OZZQAS	circumsporozoite p
23	135	7.0	401	1 OZZQAC	circumsporozoite p
24	133	6.9	419	1 OZZQAM	circumsporozoite p
25	129	6.7	485	2 A60610	circumsporozoite p
26	126	6.6	264	2 A44969	circumsporozoite p
27	125	6.5	343	2 A29319	circumsporozoite p
28	125	6.5	367	2 A32068	circumsporozoite p
29	125	6.5	378	1 OZZQAV	circumsporozoite p

ALIGNMENTS

RESULT 1

A54533

circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, Thail  
C:Species: Plasmodium falciparum  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jun-2000  
C:Accession: A54533

R.del Portillo, H.A.; Nussenzweig, R.S.; Enea, V.

Mol. Biochem. Parasitol. 24, 289-294, 1987  
A:Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.  
A:Reference number: A54533; MUID:87315205

A:Accession: A54533

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-424 <DEL>

A:Cross-references: GB:M19752; NID:g160216; PIDN:AAA29555.1; PID:g160217

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:348-402/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 11.3%; Score 216.5; DB 2; Length 424;  
Best Local Similarity 61.1%; Pred. No. 3.3e-08;  
Matches 44; Conservative 3; Mismatches 12; Indels 13; Gaps 1;

QY 26 HHHKHKLKOPGDGNPSCSVTCG-----KPKOELDYENDIEKKICKMEKC 72

DB 342 HIEQYLKKIQLSLSTEWSPSCSVTCGNGIQVRIKPGSANKPKDELVDYENDIEKKICKMEKC 401

QY 73 SSVFNVVNSNG 84

DB 402 SSVFNVVNSSIG 413

RESULT 2

OZZQAF

circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate INTM22)

C:Species: Plasmodium falciparum

C:Date: 15-Nov-1984 #sequence\_revision 15-Nov-1984 #text\_change 09-Jun-2000

C:Accession: A03388

R.Dame, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W  
Science 225, 593-599, 1984

A:Title: Structure of the gene encoding the immunodominant surface antigen on the spo

A:Reference number: A03388; MUID:84250215

A:Accession: A03388

A:Molecule type: DNA

A:Residues: 1-412 <DAM>

A:Cross-references: GB:K02194; NID:g160160; PIDN:AAA29524.1; PID:g160161

A:Experimental source: clone 7G8

C:Comment: Residues 1-16 are the probable signal sequence.

C:Comment: There are 41 copies of a 4-residue repeating unit in the middle domain of

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:336-390/Domain: thrombospondin type 1 repeat homology <THR1>

```

Query Match      11.2%; Score 215.5; DB 1; Length 412;
Best Local Similarity 62.2%; Pred. No. 3.8e-08;
Matches 46; Conservative 2; Mismatches 7; Indels 19; Gaps 3;

QY 29 KH-----KKLKQPGDGNPSPCVTCG-----KPKDELVDYNDIEKKICKME 70
Db 329 KHIOYLLKKIKN-SISTEWSFCSVTCGNGIQVRIKPGSANKPRDELVDYNDIEKKICKME 387

QY 71 KCSSVFNVNNSG 84
Db 388 KCSSVFNVNNSG 401

RESULT 3
A:Species: Plasmodium falciparum
C:Species: Plasmodium falciparum (strain wellcome)
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54529
R:Lockyer, M.J.; Schwarz, R.T.
Mol. Biochem. Parasitol. 22, 101-108, 1987
A:Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum.
A:Reference number: A54529; MUID:87115616
A:Accession: A54529
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-442 <LOC>
A:CROSS-references: GB:M15505; NID:g160214; PIDN:AAA29554.1; PID:g160215
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: tandem repeat
F:366-420/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match      11.1%; Score 213.5; DB 2; Length 442;
Best Local Similarity 59.7%; Pred. No. 5.8e-08;
Matches 43; Conservative 4; Mismatches 12; Indels 13; Gaps 1;

QY 26 HHKHKHLKQPGDGNPSPCVTCG-----KPKDELVDYNDIEKKICKMEKC 72
Db 360 HIEQYLLKTONSLSTWSPCVTCGNGIQVRIKPGSADKPKDQLDYNDIEKKICKMEKC 419

QY 73 SSVFNVNNSG 84
Db 420 SSVFNVNNSG 431

RESULT 4
A:Species: Plasmodium falciparum
C:Species: Plasmodium falciparum
C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999
C:Accession: A39756
R:Jal, A.A.; Goldman, I.F.
J. Biol. Chem. 266, 6686-6689, 1991
A:Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria
A:Reference number: A39756; MUID:91201303
A:Accession: A39756
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <LAL>
A:CROSS-references: GB:M60972; NID:g160228; PIDN:AAA29561.1; PID:g160229
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:312-366/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match      10.8%; Score 208.5; DB 2; Length 388;
Best Local Similarity 58.3%; Pred. No. 1.1e-07;
Matches 42; Conservative 4; Mismatches 13; Indels 13; Gaps 1;

QY 26 HHKHKHLKQPGDGNPSPCVTCG-----GKPKDELVDYNDIEKKICKMEKC 72
Db 306 HIEFLKIQNNLSTWSPCVTCGNGIQVRIKPGSAGKPKDQLDYNDIEKKICKMEKC 365

Query Match      10.8%; Score 207.5; DB 2; Length 405;
Best Local Similarity 51.8%; Pred. No. 1.4e-07;
Matches 43; Conservative 4; Mismatches 9; Indels 27; Gaps 2;

QY 29 KHKLKQPGD-----GNPSPCVTCG-----KPKDELVDYND 61
Db 312 KNNNEEPSDKHIKEYLNKTONSLSTWSPCVTCGNGIQVRIKPGSANKPRDELVDYND 371

QY 62 IEKKICKMEKSSVFNVNNSG 84
Db 372 IEKKICKMEKSSVFNVNNSG 394

RESULT 6
A:Species: Plasmodium falciparum
C:Species: Plasmodium falciparum
C:Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jun-2000
C:Accession: S47282
R:Toile, R.; Bujard, H.; Cooper, J.A.
submitted to the EMBL Data Library, July 1994
A:Description: Plasmodium falciparum: recombination within the C-terminal region of m
A:Reference number: S47282
A:Accession: S47282
A:Molecule type: DNA
A:Residues: 1-651 <TOL>
A:CROSS-references: EMBL:Z35329; NID:g535257; PIDN:CAA84558.1; PID:g535258
A:Experimental source: strain RO-71
C:Superfamily: major merozoite surface antigen
C:Keywords: glycoprotein; merozoite; surface antigen

```



A;Residues: 1504-1639 <MYL2>  
A;Cross-references: EMBL:X15063  
C;Superfamily: major merozoite surface antigen  
C;Keywords: glycoprotein; merozoite; surface antigen  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-1639/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 10.2%; Score 196.5; DB 2; Length 1639;  
Best Local Similarity 44.1%; Pred. No. 4.9e-06;  
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;

Qy 61 DIEKICKMEKSSVFNVNSGCFRHLDERECKLL----- 99  
Db 1526 NISQHCVKQCP-----QNSGCFRHLDERECKLLNYKQBGDKCVENPNTCNENN 1578

Qy 100 -----EDSGSNKKITCECTKPDSPKI 121  
Db 1579 GGCADAKCTEEDSGSNKKITCECTKPDSPYL 1611

RESULT 11  
SAZQK1  
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (st  
C;Species: Plasmodium falciparum  
C;Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 09-Jun-2000  
C;Accession: A25120  
R;Mackay, M.; Goman, M.; Bone, N.; Hyde, J.E.; Scaife, J.; Certia, U.; Stunnenberg, H.; B  
EMBO J. 4, 3823-3829, 1985  
A;Title: Polymorphism of the precursor for the major surface antigens of Plasmodium fal  
A;Reference number: A91040; MUID:86136024  
A;Accession: A25120  
A;Molecule type: DNA  
A;Residues: 1-1631 <MAC>  
C;Comment: The merozoite stages of different strains have strain-specific surface antigen  
C;Comment: P. falciparum has three stages: sporozoite, merozoite, and gametocyte. The m  
C;Superfamily: major merozoite surface antigen  
C;Keywords: glycoprotein; merozoite; surface antigen; tandem repeat; transmembrane prote  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-1631/Product: major merozoite surface antigen #status predicted <MAT>  
F;67-84/Region: 3-residue repeats (S-G-T/P)  
F;1614-1631/Domain: membrane anchor #status predicted <MBN>  
F;97,259,755,759,835,911,955,1049,1156,1165,1436,1563/Binding site: carbohydrate (Asn) (

Query Match 10.2%; Score 196; DB 1; Length 1631;  
Best Local Similarity 50.0%; Pred. No. 5.2e-06;  
Matches 38; Conservative 3; Mismatches 3; Indels 32; Gaps 1;

Qy 78 VVNSNGCFRHLDERECKLL-----EDSGSN 105  
Db 1528 MIQNSGCFRHLDERECKLLNYKQBGSKCVENPNTCNENNNGCDADAKCTEEDSGSN 1587

Qy 106 GKKITCECTKPDSPKI 121  
Db 1588 GKKITCECTKPDSPYL 1603

RESULT 12  
A54498  
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (is  
C;Species: Plasmodium falciparum  
C;Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jun-2000  
C;Accession: A54498  
R;Peterson, M.G.; Coppel, R.L.; McIntyre, P.; Langford, C.J.; Woodrow, G.; Brown, G.V.;  
Mol. Biochem. Parasitol. 27, 291-302, 1988  
A;Title: Variation in the precursor to the major merozoite surface antigens of Plasmodiu  
A;Reference number: A54498; MUID:88142999  
A;Accession: A54498  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1701 <PET>  
A;Cross-references: GB:M19143; NID:g160412; PID:AAA29653.1; PID:g160413

C;Superfamily: major merozoite surface antigen  
C;Keywords: surface antigen

Query Match 9.8%; Score 187.5; DB 2; Length 1701;  
Best Local Similarity 41.3%; Pred. No. 2.3e-05;  
Matches 43; Conservative 7; Mismatches 11; Indels 43; Gaps 3;

Qy 50 GKPKDELVDYNDIEKKICKMEKSSVFNVNSGCFRHLDERECKLL----- 99  
Db 1581 GKQDML-----NISQHCVKQCP-----ENSGCFRHLDERECKLLNYKQBGDKCV 1629

Qy 100 -----EDSGSNKKITCECTKPDSPKI 121  
Db 1630 ENPNTCNENNGCDADATCTEEDSGSNKKITCECTKPDSPYL 1673

## RESULT 13

A26868  
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)  
C;Species: Plasmodium falciparum  
C;Date: 19-Nov-1988 #sequence\_revision 19-Nov-1988 #text\_change 09-Jun-2000  
C;Accession: A26868  
R;Tanabe, K.; Mackay, M.; Goman, M.; Scaife, J.G.  
J. Mol. Biol. 195, 273-287, 1987

A;Title: Allelic dimorphism in a surface antigen gene of the malaria parasite Plasmod  
A;Reference number: A26868; MUID:88011243  
A;Accession: A26868  
A;Molecule type: DNA

A;Residues: 1-1701 <TAN>  
C;Superfamily: major merozoite surface antigen  
C;Keywords: surface antigen

F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-1701/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 9.8%; Score 187.5; DB 2; Length 1701;  
Best Local Similarity 41.3%; Pred. No. 2.3e-05;  
Matches 43; Conservative 7; Mismatches 11; Indels 43; Gaps 3;

Qy 50 GKPKDELVDYNDIEKKICKMEKSSVFNVNSGCFRHLDERECKLL----- 99  
Db 1581 GKQDML-----NISQHCVKQCP-----ENSGCFRHLDERECKLLNYKQBGDKCE 1629

Qy 100 -----EDSGSNKKITCECTKPDSPKI 121  
Db 1630 ENPNTCNENNGCDADATCTEEDSGSNKKITCECTKPDSPYL 1673

## RESULT 14

A24594  
probable major surface antigen (83K, 19K, 42K) precursor - malaria parasite (Plasmodi  
C;Species: Plasmodium falciparum  
C;Date: 29-Aug-1987 #sequence\_revision 29-Aug-1987 #text\_change 09-Jun-2000  
C;Accession: A24594  
R;Holder, A.A.; Lockyer, M.J.; Odink, K.G.; Sandhu, J.S.; Riveros-Moreno, V.; Nicholl  
Nature 317, 270-273, 1985

A;Title: Primary structure of the precursor to the three major surface antigens of Pl  
A;Reference number: A24594; MUID:86014355  
A;Accession: A24594  
A;Molecule type: DNA  
A;Residues: 1-1640 <HOL>

C;Superfamily: major merozoite surface antigen  
C;Keywords: surface antigen

Query Match 9.7%; Score 186; DB 2; Length 1640;  
Best Local Similarity 47.4%; Pred. No. 2.8e-05;  
Matches 36; Conservative 4; Mismatches 4; Indels 32; Gaps 1;

Qy 78 VVNSNGCFRHLDERECKLL-----EDSGSN 105  
Db 1537 MIQSSGCFRHLDERECKLLNYKQBGSKCVENPNTCNENNGCDADAKCTEEDSGSN 1596





**THIS PAGE BLANK (USPTO)**



```
QY 73 SSVFNVNNSG 84
DB 402 SSVFNVNNSG 413
|||||||: |

RESULT 2
CSP_PLAFW STANDARD; PRT; 412 AA.
ID CSP_PLAFW
AC P02893;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Rasmussen G.S., Reddy E.P., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum."
RL Science 225:593-599(1984).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M15505; AAA29554.1; -
CC PIR; A54529; A54529.
CC InterPro; IPR003067; Crcmsprzoite.
CC Pfam; PF00090; tsp_1; 1.
CC PRINTS; PR01303; CRCMSPRZOITE.
CC SMART; SM00209; TSP1; 1.
CC Malaria; Sporozoite; Repeat; Signal.
KW SIGNAL 1 16
FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 230 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
FT SEQUENCE 412 AA; 44420 MW; 1EEED3DE90965F8 CRC64;
SQ
Query Match 11.2%; Score 215.5; DB 1; Length 412;
Best Local Similarity 62.2%; Pred. No. 5.9e-08;
Matches 46; Conservative 2; Mismatches 7; Indels 19; Gaps 3;
QY 29 KH-----KKLKQDGNPNWSPCVTCG-----KPKDELDYNDIEKKICKME 70
DB 329 KHIEQYLKKIKN-SISTWSPCVTCGNGIQVRIKPGSADKPKDQDLYNDIEKKICKME 387
QY 71 KCSSVFNNSG 84
DB 388 KCSSVFNNSG 401
|||||||: |

RESULT 3
CSP_PLAFW STANDARD; PRT; 442 AA.
ID CSP_PLAFW
AC P25694;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91201303; PubMed=2016283;
RX
```

RA Lal A.A., Goldman I.F.;

RT "Circumsporozoite protein gene from Plasmodium reichenowi, a

RT chimpanzee malaria parasite evolutionarily related to the human

RT malaria parasite Plasmodium falciparum.";

RL J. Biol. Chem. 266:6686-6689(1991).

CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT

CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE

CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE

CC VERTEBRATE HOST).

CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR

CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES

CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.

CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M60972; AAA29561.1; -

DR PIR: A39756; A39756.

DR InterPro: IPR003067; Crcmsprzoite.

DR InterPro: IPR000884; TSPI.

DR Pfam: PF00090; tsp\_1; 1.

DR PRINTS: PR01303; CRCMSPRZOITE.

DR SMART: SM00209; TSPI; 1.

KW Malaria: Sporozoite; Repeat; Signal.

FT SIGNAL 1 16 PROBABLE.

FT CHAIN 17 388

FT DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.

SQ SEQUENCE 388 AA; 42245 MW; C031EEFBEZE35604 CRC64;

Query Match 10.8%; Score 208.5; DB 1; Length 388;

Best Local Similarity 58.3%; Pred. No. 1.7e-07;

Matches 42; Conservative 4; Mismatches 13; Indels 13; Gaps 1;

OY 26 HHKHKHKLKPGDGNPWSVSCVTC-----GKPKDELDYENDIEKKICKMEKC 72

DB 306 HIEFLUKIQNNLSTWSPSCVTCGNGIQVRIKPGSAGKPKDQLDYENDLEKKICKMEKC 365

OY 73 SSVFNVNNSG 84

DB 366 SSVFNVNNSG 377

RESULT 5

CSP\_PLAFO STANDARD; PRT; 397 AA.

AC P19597; Q25798;

DT 01-FEB-1991 (Rel. 17, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).

OS Plasmodium falciparum (isolate NF54).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=5843;

RN [1]

RN SEQUENCE FROM N.A.

RX MEDLINE=89345189; PubMed=2668895;

RA Campbell J.R.;

RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria

RT candidate vaccine antigen.";

RL Nucleic Acids Res. 17:5854-5854(1989).

RN [2]

RN REVISIONS.

RP Campbell J.R.;

RA Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=92155298; PubMed=1346766;

RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,

RA Thomas A.W., Baqar S., Cochran M.A., Thanassi J., Levine M.M.,

RA Hackett C.S.;

RT "Plasmodium falciparum: in vitro characterization and human

RT infectivity of a cloned line.";

RL Exp. Parasitol. 74:159-168(1992).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=89364998; PubMed=2671723;

RA Caspers P., Gentz R., Matile H., Pink J.R., Sinigaglia F.;

RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum

RT isolate used in malaria vaccine trials.";

RL Mol. Biochem. Parasitol. 35:185-190(1989).

CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT

CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE

CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE

CC VERTEBRATE HOST).

CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR

CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES

CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.

CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X15363; CAA33421.1; -

DR EMBL: M83886; AAA29521.1; -

DR EMBL: M22982; AAA29527.1; -

DR PIR: S05428; S05428.

DR PIR: A45527; A45527.

DR InterPro: IPR003067; Crcmsprzoite.

DR InterPro: IPR000884; TSPI.

DR Pfam: PF00090; tsp\_1; 1.

DR PRINTS: PR01303; CRCMSPRZOITE.

DR SMART: SM00209; TSPI; 1.

KW Malaria: Sporozoite; Repeat; Signal.

FT SIGNAL 1 16 PROBABLE.

FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.

FT DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.

FT CONFLICT 194 194 A -> ANPNAPNA (IN REF. 4).

SQ SEQUENCE 397 AA; 42646 MW; 9E81146F59EBCEA3 CRC64;

Query Match 10.8%; Score 207.5; DB 1; Length 397;

Best Local Similarity 51.8%; Pred. No. 2e-07;

Matches 43; Conservative 4; Mismatches 9; Indels 27; Gaps 2;

OY 29 KHKKLKQPGD-----GNPWSVSCVTCG-----KPKDELDYEND 61

DB 304 KNNNEEPSDKHIKYLKINQSLSTWSPSCVTCGNGIQVRIKPGSANKPKDLDYAND 363

OY 62 IEKKICKMEKCSVFNVNNSG 84

DB 364 IEKKICKMEKCSVFNVNNSG 386

RESULT 6

MSPI\_PLAF3

ID MSPI\_PLAF3 STANDARD; PRT; 1682 AA.

AC P19598; Q25921;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)

DE (PMMSA) (P190).

GN MSP-1.

OS Plasmodium falciparum (isolate ro-33 / Ghana).

01-WAR-1989 (Rel. 10, Last sequence update)  
01-OCT-1996 (Rel. 34, Last annotation update)  
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)  
DE (PMMSA) (P195).  
GN MSP-1.  
OS Plasmodium falciparum (isolate Camp / Malaysia).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5835;  
RN [1]  
RN SEQUENCE OF 1-1103 FROM N.A.  
RN MEDLINE=86205236; PubMed=3517809;  
RA Weber J.L., Leininger W.M., Lyon J.A.;  
RT "Variation in the gene encoding a major merozoite surface antigen of  
RT the human malaria parasite Plasmodium falciparum.";  
RL Nucleic Acids Res. 14:3311-3323(1986).  
RN [2]  
RN SEQUENCE OF 1104-1726 FROM N.A.  
RN MEDLINE=88143999; PubMed=3278296;  
RA Weber J.L., Sim B.K.L., Lyon J.A., Wolff R.;  
RT "Merozoite surface protein sequence from the Camp strain of the human  
RT malaria parasite Plasmodium falciparum.";  
RL Nucleic Acids Res. 16:1206-1206(1988).  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR  
CC (POTENTIAL).  
CC -1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42  
CC KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF  
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/annouce/>)  
CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
CC -----  
CC EMBL; X03831; CAA27446.1; --  
DR PTR; A23386; SAZQGM.  
DR InterPro; IPR000561; EGF-like.  
DR Pfam; PF00008; EGF; 1.  
KW Malaria; Merozoite; Polypotein; Repeat; Signal; Glycoprotein;  
KW Transmembrane; GPI-anch.  
KW SIGNAL 1 19  
FT CHAIN 20 1726 POTENTIAL. MEROZOITE SURFACE PROTEIN 1.  
FT FT CARBOHYD 133 133 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 501 501 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 567 567 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 638 638 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 827 827 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 924 924 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 944 944 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 990 990 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 1221 1221 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT CARBOHYD 1613 1613 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 1726 AA; 196197 MW; DDB8AD45FA352BCF3 CRC64;  
Query Match 10.4%; Score 200.5; DB 1; Length 1726;  
Best Local Similarity 43.3%; Pred. No. 3.5e-06;  
Mechan. 45; Conservative 6; Indels 43; Gaps 3

```

Query Match      10.48; Score 200.5; DB 1; Length 1682;
Best Local Similarity 43.3%; Pred. NO. 3.4e-06;
Matches 45; Conservative 6; Mismatches 10; Indels 43; Gaps 3;
QY 50 GKPKDELDYENDIEKKICKMEKCSVFNVNNSGCCFRLHDERECKLL-----99
Db 1562 GKFDML----NISQHCVKKQCP-----QNSGCFRHLDERECKLLNYKQEGDKCV 1610
QY 100 -----EDSGNGKKITCECTKPDSKPI 121
Db 1611 ENPNPTCENNCGDADAKATEEDSGSNGKKITCECTKPDSYPL 1654

RESULT 7
MSPI_PLAFC
ID MSPI_PLAFC STANDARD; PRT; 1726 AA.
AC P04934;
DT 13-AUG-1987 (Rel. 05, Created)

```

```

RESULT      8
MSPL_PLAFF STANDARD;          PRT: 1726 AA.
ID   MSPI_PLAFF
P50495;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
(PMMSA) (GPI195).
GN   MSP-1.
OS   Plasmodium falciparum (isolate Palo Alto / Uganda).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=57270;
RN   [1]
RS   SEQUENCE FROM N.A.
RX   MEDLINE=89005525; PubMed=3049134;
RA   Chang S.P., Kramer K.J., Yamaga K.M., Kato A., Case S.E.,
RA   Siddiqui W.A.;
RT   "Plasmodium falciparum: gene structure and hydrophathy profile of the
RT   major merozoite surface antigen (gpi95) of the Uganda-Palo Alto
RT   isolate.";
RL   Exp. Parasitol. 67:11-11(1988).
CC   -I- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC   -I- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
-----
CC   THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@lsb-sib.ch).
-----
DR   EMBL; M37213; AAA29611.1; -.
DR   InterPro: IPR000561; EGF-like.
DR   Pfam; PF00008; EGF; 1.
KW   Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
Transmembrane; GPI-anchor.
FT   SIGNAL        1..19      POTENTIAL.
FT   CHAIN         20..1726   MEROZOITE SURFACE PROTEIN 1.
FT   CARBOHYD     133..133    N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD     272..272    N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD     501..501    N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD     567..567    N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD     638..638    N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD     827..827    N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD     924..924    N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD     944..944    N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD     990..990    N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD    1016..1016   N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD    1114..1114   N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD    1221..1221   N-LINKED (GLCNAC. .) (POTENTIAL).
FT   CARBOHYD    1613..1613   N-LINKED (GLCNAC. .) (POTENTIAL).
SQ   SEQUENCE      1726 AA; 196174 MW; 5859CEFA2F9A026 CRC64;
Query Match      10.4%; Score 200.5; DB 1; Length 1726;
Best Local Similarity 43.3%; Pred. No. 3.5e-06;
Matches 45; Conservative 6; Mismatches 10; Indels 43; Gaps 3;
QY   50 GKPKDELVDYNDIEKKICKMEKCSSVFNVNSNGCFRHLDERECKCLL----- 99
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB   1606 GKFDML-----NISQHQCVRKQP-----ENSGCFRHLDERECKCLLNKGDKCV 1654
QY   100 -----EDSGNKKKITCECTKPDSPKI 121
||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB   1655 ENPNPTCENNNGCGDAKCTEEDSGSNKKKITCECTKPDSPYL 1698

```

ID	MSPI_PLAFLK	STANDARD;	PRT; 1630 AA.
AC	P04932;		
DT	13-AUG-1987 (Rel. 05, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)		
DT	01-OCT-1996 (Rel. 34, Last annotation update)		
DE	MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS) (PMMSA) (P190).		
DE	MSP-1.		
OS	Plasmodium falciparum (isolate K1 / Thailand).		
OX	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=5839;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86136024; PubMed=3004972;		
RA	Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,		
RA	Stunnenberg H., Bujard H.;		
RT	"Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum merozoites: studies at the genetic level.";		
RL	EMBO J. 4:3823-3829(1985).		
RN	[2]		
RP	REVISIONS, SEQUENCE FROM N.A.		
RP	Pan W., Tolle R., Bujard H.;		
RA	Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases..		
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (POTENTIAL).		
CC	-1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	EMBL: X03371; CAA27070.1; -.		
DR	PIR: A25120; SAZOK1.		
DR	InterPro: IPR000561; EGF-like.		
DR	Fam; PF00008; EGF; 1.		
KW	Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein; Transmembrane; GPI-anchor.		
KW	SIGNAL	1 19	POTENTIAL.
FT	CHAIN	20 1630	MEROZOITE SURFACE PROTEIN 1.
FT	DOMAIN	67 84	TRIPEPTIDE SG(TP) REPEAT.
FT	TRANSMEM	1614 1630	MEMBRANE ANCHOR.
FT	CARBOHYD	97 97	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	259 259	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	755 755	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	759 759	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	774 774	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	835 835	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	911 911	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	955 955	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1049 1049	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1156 1156	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1165 1165	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1436 1436	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1517 1517	N-LINKED (GLCNAC. .) (POTENTIAL).
SO	SEQUENCE	1630 AA; 187289 MW; ABDEG3CE0A46322 CRC64;	

Query Match 10.2%; Score 196.5; DB 1; Length 1630;  
 Best Local Similarity 44.1%; Pred. No. 6.le-06;  
 Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;

QY 61 DIEKKICKMEKCSSVFNVNNSSCFRRHLDERECKLL----- 99  
 :| : | ::| |||||  
 Db 1517 NISQHOCVKKQCP-----QNSGCFRRHLDERECKLLNYKGDKVCNPNTCENN 1569

OY 100 -----EDSGSNKGKITCCTCRPSKPI 121

```
Db 1570 GGCDAKTEEDSGNGKKITCECTKPDSPYPL 1602
|||||
RESULT 10
MSPL_PLAFW STANDARD; PRT; 1639 AA.
AC P04933;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA) (P195).
GN MSP-1.
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=6014355; PubMed=2995820;
RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
RA Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
RA Freeman R.R.;
RT "Primary structure of the precursor to the three major surface
RT antigens of Plasmodium falciparum merozoites."
RL Nature 317:270-273(1985).
RN [2]
RP REVISIONS.
RA Holder A.A.;
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X02919; CAA26676.1; -.
DR PIR; A24594; A24594.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
DR Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchored.
FT SIGNAL 1 19
FT CHAIN 20 1639
FT CARBOHYD 116 116 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 764 764 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 768 768 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 783 783 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 844 844 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 920 920 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;

Query Match 10.2%; Score 196.5; DB 1; Length 1639;
Best Local Similarity 44.1%; Pred. No. 6.1e-06;
Matches 41; Conservative 5; Mismatches 8; Indels 39; Gaps 2;

QY 61,PIEKICKMEKCSVFNVNNSGCFRHLDERECKCLL----- 99
|||||
```

```
Db 1526 NISQHCQVKQCP-----QNSGCFRHLDERECKCLLNTYKQEGDKCVENPNTCMENN 1578
|||||
QY 100 -----EDSGNGKKITCECTKPDSPKI 121
|||||
Db 1579 GGCDAKTEEDSGNGKKITCECTKPDSPYPL 1611
|||||
RESULT 11
MSPL_PLAFF STANDARD; PRT; 1701 AA.
AC P13819;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
DE (PMMSA).
GN MSP-1.
OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5837;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88142999; PubMed=2449612;
RA Peterson M.G., Coppel R.L., McIntyre P., Langford C.J., Woodrow G.,
RA Brown G.V., Anders R.F., Kemp D.J.;
RT "Variation in the precursor to the major merozoite surface antigens
RT of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 27:291-302(1988).
RN [2]
RP SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(POTENTIAL).
CC -!- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 42
CC KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M19143; AAA29653.1; -.
DR PIR; A54498; A54498.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
DR Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchored.
FT SIGNAL 1 19
FT CHAIN 20 1701
FT CARBOHYD 110 110 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 470 470 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 802 802 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 919 919 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 991 991 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1196 1196 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1588 1588 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1701 AA; 193719 MW; 3920B75E73D38552 CRC64;

Query Match 9.8%; Score 187.5; DB 1; Length 1701;
Best Local Similarity 41.3%; Pred. No. 2.7e-05;
Matches 43; Conservative 7; Mismatches 11; Indels 43; Gaps 3;

QY 50 GKPRDELDYNDIEKKICKMEKCSVFNVNNSGCFRHLDERECKCLL----- 99
|||||
```







Search completed: January 29, 2002, 11:13:36  
Job time: 812 sec

**THIS PAGE BLANK (USPRO)**





```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-760-797A-3

Query Match 10.8%; Score 207.5; DB 2; Length 424;
Best Local Similarity 51.8%; Pred. No. 1.9e-11;
Matches 43; Conservative 4; Mismatches 9; Indels 27; Gaps 2;

QY 29 KHKLLKQPGD-----GNPWSPCSVTCG-----KPKDELVDYND 61
   | : : : |
   | : : : |
Db 110 KNNNEEPSDKHIKEYLNKIQNSLSTWSPSVTCGNGIQVRIKPGSANKPKDELVDYND 169
   | : : : |
   | : : : |
QY 62 IEKKICKMEKCSVFNVNNSG 84
   | : : : |
   | : : : |
Db 170 IEKKICKMEKCSVFNVNNSG 192
   | : : : |
   | : : : |

RESULT 5
US-08-932-929B-3
; Sequence 3, Application US/08932929B
; Patent No. 6169171
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBsAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,929B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/760,797
; FILING DATE: 04-DEC-1996
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
```

```

; REFERENCE/DOCKET NUMBER: B45015-1FWC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-932-929B-3

Query Match 10.8%; Score 207.5; DB 4; Length 424;
Best Local Similarity 51.8%; Pred. No. 1.9e-11;
Matches 43; Conservative 4; Mismatches 9; Indels 27; Gaps 2;

QY 29 KHKLLKQPGD-----GNPWSPCSVTCG-----KPKDELVDYND 61
   | : : : |
   | : : : |
Db 110 KNNNEEPSDKHIKEYLNKIQNSLSTWSPSVTCGNGIQVRIKPGSANKPKDELVDYND 169
   | : : : |
   | : : : |
QY 62 IEKKICKMEKCSVFNVNNSG 84
   | : : : |
   | : : : |
Db 170 IEKKICKMEKCSVFNVNNSG 192
   | : : : |
   | : : : |

RESULT 6
US-08-257-073-9
; Sequence 9, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: de Taisne, Charles
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 628 amino acids
```

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-257-073-9

Query Match          9.18; Score 175.5; DB 1; Length 628;
Best Local Similarity 25.08; Pred. No. 3.5e-08;
Matches 82; Conservative 24; Mismatches 89; Indels 133; Gaps 13;

Qy 22 DHHHHHHKHKLLKPGDGNPWSPCSVTCGKPKDELVDLEKKIC---KMEKSSVFNV 78
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 196 DEMRHFYKDNKYVNLDD--ELTLCRHAGNMIIPDNKNSNYKYPAVYDDKDKCHILYIA 253
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 79 VNSNSG---CFRHLDEREECKLLEDGSSNKKITCECTKPDSPITVOYDNF-NANPNAN 134
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 254 AQENNGPRYCNKDESKRSMCEFGA-RPAKDISFQNYTILSKNVV--DNWEKVCPRKN 306
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 135 -PNAN---PGNCEDIDPHNEFSALDIGNAEKYDKMDPEQHGKSLTPLEELYKPNDKS 189
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 307 LONAKFGLWVGNCEDIDPHVNEFFPAIDLFECNKL----- 340
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 190 LYQYIKANSKFITGIELSNTFTINNAGQGHMHGNEREDERTLTKEYEDIVLKEFTYMINF 249
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 341 -----VFELS----- 345
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 250 GRGQNYWEHPYQKSDQPKQYEOHLYEYKIKGKPLDKFGNIYDVHYEHSSPSSTKSSSP 309
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 346 -----ASDPKQYEOHLYDYBKIKGG-----FNKNASMIKS--- 377
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 310 SNVKSASLATRLMKFKFAIRDFFGISY 337
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 378 RRIKSAFLPT---GAFKADRYKSHGKY 402
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 7
US-08-290-919-3
; Sequence 3, Application US/08290919
; Patent No. 5720959
; GENERAL INFORMATION:
; APPLICANT: HOLDER, ANTHONY A.
; APPLICANT: BLACKMAN, MICHAEL J.
; APPLICANT: CHAPPEL, JONATHAN A.
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
; TITLE OF INVENTION: VACCINE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,919
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9203821.5
; FILING DATE: 22-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00367
; FILING DATE: 22-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16 773

```



```

Query Match          6.7%  Score 128.5;  DB 4;  Length 66;
Best Local Similarity 51.8%;  Pred. NO. 3.4e-05;
Matches 29;  Conservative 3;  Mismatches 11;  Indels 13;  Caps 1;

Qy 31 K L K Q P G D G N P -----W S C S V T C G K P K D E L Y E N D I E K K I C K M E K C S 73
      |||  ||  ||  :  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 11 K T K N S I S C N P N P N A N P N A N P N A N P N A N P K P K D E L Y E N D I E K K I C K M E K C S 66

```

```

RESULT 10
US-08-060-988A-35
; Sequence 35, Application US/08060988A
; Patent No. 6294322
; GENERAL INFORMATION:
; APPLICANT: Berzofsky, Jay A.
; APPLICANT: Ahlers, Jeffrey D.

```

```

APPLICANT: Nairai, Peter
TITLE OF INVENTION: MULTIDETERMINANT PEPTIDES
TITLE OF INVENTION: THAT ELICIT
TITLE OF INVENTION: HELPER T-LYMPHOCYTE, CYTOTOXIC T LYMPHOCYTE AND
TITLE OF INVENTION: NEUTRALIZING ANTIBODY RESPONSES AGAINST HIV-1
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/060,988A
FILING DATE: 14-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,311
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/751,998
FILING DATE: 29-AUG-1991
APPLICATION NUMBER: 07/148,692
FILING DATE: 26-JAN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Beattie, Ingrid A.
REGISTRATION NUMBER: P-42,306
REFERENCE/DOCKET NUMBER: 08830/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-060-988A-35

Query Match 6.7% Score 128.5; DB 4; Length 66;
Best Local Similarity 51.8%; Pred. No. 3.4e-05;
Matches 29; Conservative 3; Mismatches 11; Indels 13; Gaps 1;

QY 31 KLLKQPQDGNP-----WSPCSVTCGPKDELVDYENDIEKKICKMEKCS 73
||| || : |||||
Db 11 KTKINSTSCNPANPNANPNANPNANPNANPNANPNAPKDELVDYENDIEKKICKMEKCS 66

```

Mon Feb 4 15:23:41 2002

```
RESULT 11
PCT-US94-05142-35
; Sequence 35, Application PC/TUS9405142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITE SYNTHETIC PEPTIDE CONSTRUCT
; TITLE OF INVENTION: ELICITING NEUTRALIZING ANTIBODIES AND CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTES AGAINST HIV
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05142
; FILING DATE: 13-MAY-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/060,988
; FILING DATE: 14-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30330
; REFERENCE/DOCKET NUMBER: 1173-434P
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 66 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..66
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "peptide from P. falciparum CS antigen"
PCT-US94-05142-35

Query Match 6.7%; Score 128.5; DB 5; Length 66;
Best Local Similarity 51.8%; Pred. No. 3.4e-05;
Matches 29; Conservative 3; Mismatches 11; Indels 13; Gaps 1;

QY 31 KKLKQPGDGNP-----WSPCSTGCGKPKDLDYENDIEKICKMEKCS 73
Db 11 KKAIKSICNPANPNANPNANPNANPNANPNANPNNAKPKDLDYENDIEKICKMEKCS 66

RESULT 12
5171843-9
; Patent No. 5171843
; APPLICANT: NUSSENZWEIG, VICTOR
; TITLE OF INVENTION: IMMUNOGENIC POLYPEPTIDE AND METHOD FOR
; PURIFYING IT
; NUMBER OF SEQUENCES: 13
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/175,112
; FILING DATE: 30-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 754,645
; FILING DATE: 9-JUL-1985

us-09-763-397a-2.rai

; APPLICATION NUMBER: 115,634
; FILING DATE: 26-OCT-1987
; APPLICATION NUMBER: 649,903
; FILING DATE: 12-SEP-1984
; SEQ ID NO:9:
; LENGTH: 378
5171843-9

Query Match 6.5%; Score 125; DB 6; Length 378;
Best Local Similarity 40.0%; Pred. No. 0.00093;
Matches 24; Conservative 10; Mismatches 10; Indels 16; Gaps 2;

QY 39 GNPWSPCSTGCG-----RPKDELVDYENDIEKICKMEKCSVFNVNNSG 84
Db 310 GTETPCSTGCGVGRVRSVNAANKPEDLT--LNDLETDVCTMDKAGIFNVVNSLG 367

RESULT 13
US-08-313-288B-14
; Sequence 14, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 559 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-313-288B-14

Query Match 6.5%; Score 125; DB 1; Length 559;
Best Local Similarity 21.0%; Pred. No. 0.0017;
Matches 74; Conservative 46; Mismatches 104; Indels 128; Gaps 17;

QY 19 IYADHHHHHHKH-----KKLKQPGDGNPSPCSTGCGK----PKDELVDY-- 59
Db 214 LYADSAWENKVNIGPPPMKAVCEVEKTASCGVWDENSPCSTGCGKTRSRKREILHEGC 273
QY 60 -NDIEKICKMEKCSVFNVNNSGCGFRHLDERECKCLLEDGSGNGKKITCCTKPKDS 118
; FILING DATE: 9-JUL-1985
```

```
Db 274 TSEIQEQ-CEERCPPK-----WEPLDVPDE-----PEDDQPRPRGDNSSVQKPEE 318
Qy 119 KPIVOYDNFNANPNP-----NANPDG-----NCE-----DI-----PHVNEFSAIDL--- 157
Db 319 NII-----DNNQPSNPSEGGDENPFGDLDPENPNPDIPQKPNIPEDSEKEVPSD 375
Qy 158 -----GNAEKYDKWDEPOHYGKSLTLEELYKPKDKSLYQYIKANSKFIGITELSNTF 210
Db 376 VPKNPEDDREENFDLPKPKENKHDQNNL-----PNDKS-----DRNIPYSPLPPKV 422
Qy 211 INN-----ACQGHMHGNEREDERTLTKEYEDIVLKEFTYMINFGRGONYWEHPYO 261
Db 423 LNERKQSQDQSDNNGNRHVPNSDRET-----RPHG 455
Qy 262 KSDQPKQYEQHLTDYKEIKEGKPLDKFGNIYDHYEHSSPSTKSSSPSNVK 313
Db 456 RNNENRSYRKYNTPKPHPE-----EHEKPDNNKKKGESDNK 494

RESULT 14
US-08-155-888-2
; Sequence 2, Application US/08155888
; Patent No. 6066623
; GENERAL INFORMATION:
; APPLICANT: Hoffman, Stephen L.
; APPLICANT: Hedstrom, Richard C.
; APPLICANT: Sedegah, Martha
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE PROTECTIVE
; TITLE OF INVENTION: AGAINST MALARIA, METHODS OF PROTECTION AND VECTOR FOR
; TITLE OF INVENTION: DELIVERING POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Naval Medical Res. & Dev. Cmd.
; STREET: Bldg. 1, T-12 8901 Wisconsin Ave.
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/155.888
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spevack, A. David
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: N.C. 75,851
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 295-6759
; TELEFAX: (202) 295-1022
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-155-888-2

Query Match 6.5%; Score 124.5; DB 3; Length 478;
Best Local Similarity 45.5%; Pred. No. 0.0015;
Matches 25; Conservative 9; Mismatches 8; Indels 13; Gaps 2;

Qy 42 WSPCVTCG-----KPKDELVDYNDIEKKCKCKCKCKSSVNVNNSG 84
Db 414 WSQCSVTCGVRVRKRVNKNQENLTLE-DIDTEICKMDKCSSIFNIVNSLG 467

RESULT 15
```

```
US-08-568-459A-4
; Sequence 4, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568.459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
US-08-568-459A-4

Query Match 6.4%; Score 124; DB 2; Length 1435;
Best Local Similarity 23.1%; Pred. No. 0.0081;
Matches 61; Conservative 36; Mismatches 113; Indels 54; Gaps 7;

Qy 44 PCSVTCGPKDELVDYNDIEKK---ICKMEKCSSVNVNNSGCFRHLDERECKCLLE 100
Db 859 PLUSDDVRPKKELEDQNSDESEETVYVNHISKSPSINNGDDSGSATVSSESSNTGLSI 918
Qy 101 DSGSNGKKITCECTKPKDPSKPIVQYDNFN-----ANPNANPNANPDG 141
Db 919 DDDRNGDTFVRTDANTEDVIRKENADKDEKGADEERHSTSELSSPSEKMLTDNEG 978
Qy 142 ----NCEDI-----PHVNEFSAIDLGAEK--YDKMDEPOHYGKSLTLEELYKPNK 188
Db 979 GNSLNHEEVKEHTSNSDNVQSGGVNMMVKEKLDLTLENSSSLDEGKAHEELSEPNLS 1038
Qy 189 SLYQYIKANSKFIGITELSNTEINNAGQGHMHGNEREDERTLTKEYEDIVLKEFTYMIN 248
Db 1039 SQDDMSNTPGLDNTSEETTERISN-----NEYKVNREDEERTLTKEYEDIVLK----- 1087
Qy 249 FGRGQNYWEHPYQKSDQPKQYEQH 272
Db 1088 -----SHMNRSDDGELYDEN 1103
```

Search completed: January 29, 2002, 10:24:00  
Job time: 506 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:59:24 ; Search time 285.36 Seconds  
(without alignments)  
179.406 Million cell updates/sec

Title: US-09-763-397A-2  
Perfect score: 1923  
Sequence: 1 MKFLVNVALFVFMVVIYSIY.....DFGSIYKVLAKYKDDLE 350

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues  
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_17:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.rodent:\*
- 12: sp.virus:\*
- 13: sp.vertebrate:\*
- 14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216.5	11.3	115	5 Q9TVW7	Q9TVW7 plasmodium
2	216.5	11.3	424	5 Q27425	Q27425 plasmodium
3	216.5	11.3	424	5 Q99256	Q99256 plasmodium
4	216.5	11.3	432	5 Q27246	Q27246 plasmodium
5	216.5	11.3	432	5 Q25827	Q25827 plasmodium
6	215.5	11.2	80	5 Q900Q4	Q900Q4 plasmodium
7	215.5	11.2	80	5 Q900P1	Q900P1 plasmodium
8	213.5	11.1	115	5 Q90934	Q90934 plasmodium
9	213.5	11.1	117	5 Q25795	Q25795 plasmodium
10	213.5	11.1	416	5 Q25829	Q25829 plasmodium
11	213.5	11.1	442	5 Q25830	Q25830 plasmodium
12	212.5	11.1	80	5 Q900P7	Q900P7 plasmodium
13	212.5	11.1	115	5 Q25837	Q25837 plasmodium
14	212.5	11.1	115	5 Q25839	Q25839 plasmodium
15	212.5	11.1	420	5 Q25831	Q25831 plasmodium
16	212.5	11.1	436	5 Q25828	Q25828 plasmodium
17	212.5	11.1	452	5 Q25834	Q25834 plasmodium
18	211.5	11.0	80	5 Q9TVP9	Q9TVP9 plasmodium
19	211.5	11.0	115	5 Q25835	Q25835 plasmodium

20	211.5	11.0	115	5 Q25836	Q25836 plasmodium
21	211.5	11.0	117	5 Q25794	Q25794 plasmodium
22	210.5	10.9	80	5 Q9U0P8	Q9U0P8 plasmodium
23	210.5	10.9	80	5 Q9TW97	Q9TW97 plasmodium
24	210.5	10.9	80	5 Q9TVN9	Q9TVN9 plasmodium
25	209.5	10.9	80	5 Q9U0Q0	Q9U0Q0 plasmodium
26	209.5	10.9	420	5 Q25838	Q25838 plasmodium
27	208.5	10.8	80	5 Q9U0P3	Q9U0P3 plasmodium
28	208.5	10.8	80	5 Q9TW00	Q9TW00 plasmodium
29	208.5	10.8	117	5 Q25797	Q25797 plasmodium
30	208.5	10.8	408	5 Q25729	Q25729 plasmodium
31	207.5	10.8	80	5 Q9U0Q3	Q9U0Q3 plasmodium
32	207.5	10.8	80	5 Q9TW01	Q9TW01 plasmodium
33	207.5	10.8	117	5 Q25796	Q25796 plasmodium
34	207.5	10.8	393	5 Q99255	Q99255 plasmodium
35	206.5	10.7	80	5 Q9U0P5	Q9U0P5 plasmodium
36	206.5	10.7	80	5 Q9TW76	Q9TW76 plasmodium
37	206.5	10.7	436	5 Q27325	Q27325 plasmodium
38	205.5	10.7	80	5 Q9U0P9	Q9U0P9 plasmodium
39	204.5	10.6	80	5 Q9U0Q1	Q9U0Q1 plasmodium
40	204.5	10.6	80	5 Q9U0P2	Q9U0P2 plasmodium
41	203.5	10.6	80	5 Q9U0P4	Q9U0P4 plasmodium
42	203.5	10.6	80	5 Q9TVQ0	Q9TVQ0 plasmodium
43	202.5	10.5	80	5 Q9U0P6	Q9U0P6 plasmodium
44	202.5	10.5	80	5 Q9TW83	Q9TW83 plasmodium
45	200.5	10.4	116	5 Q9TYE3	Q9TYE3 plasmodium

## ALIGNMENTS

RESULT 1  
Q9TVW7 PRELIMINARY; PRT; 115 AA.  
ID Q9TVW7  
AC Q9TVW7  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).  
GN CSP.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RC STRAIN=815, 947, AND 808;  
RX MEDLINE=95077069; PubMed=7985759;  
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;  
RT "Allelic variation in the circumsporozoite protein of Plasmodium  
falciparum from Thai field isolates";  
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).  
DR EMBL; M83153; AAA29566.1; -;  
DR EMBL; M83171; AAA29549.1; -;  
DR EMBL; M83151; AAA29564.1; -;  
DR InterPro; IPR000884; TSPI.  
DR InterPro; IPR003067; Crcmsprzoite.  
DR Pfam; PF00090; tsp.1; 1  
DR PRINTS; PR01303; Crcmsprzoite.  
DR SMART; SM00209; TSPI; 1.  
FT NON\_TER 1  
SQ SEQUENCE 115 AA; 12925 MW; 697DF63EBEBDA90 CRC64;

Query Match 11.3%; Score 216.5; DB 5; Length 115;  
Best Local Similarity 61.1%; Pred. No. 1.7e-09;  
Matches 44; Conservative 3; Mismatches 12; Indels 13; Gaps 1;

QY 26 HHKHKHKLKQPCDGNPWSVTCG-----KPKDELVDYENDIEKKICKMEKC 72  
Db 33 HIEQYLUKKIQLNSLSTWSPSVTCGNGIQVRIKPGSANKPKDELVDYENDIEKKICKMEKC 92  
QY 73 SSVFNVVNSNSG 84  
|||||||:|



```
DR EMBL; M83166; AAA29544.1; -.
DR EMBL; M83168; AAA29546.1; -.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
SQ SEQUENCE 432 AA; 46414 MW; 8787E6005578873A CRC64;

Query Match          11.3%; Score 216.5; DB 5; Length 432;
Best Local Similarity 61.1%; Pred. No. 7.7e-09;
Matches 44; Conservative 3; Mismatches 12; Indels 13; Gaps 1;

Qy 26 HHHKKLQPCDGNPSPCVTCG-----KPKDELDYENDIEKKICKMEKC 72
   |::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 350 HIEQYLLKIQNSLSTWSPCVTCGNGIQVRIKPGSANKPKDELDYENDIEKKICKMEKC 409

Qy 73 SSVFNVNSNG 84
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 410 SSVFNVNSSIG 421

RESULT 5
ID Q25827 PRELIMINARY; PRT; 432 AA.
AC Q25827;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=838;
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum."
RL Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=838;
RA la Cruz V.F.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=838;
RA Jongwutives S., Tanabe K., Kanbara H.;
RL Mol. Biochem. Parasitol. 0:0-0(0).
DR EMBL; M83165; AAA29543.1; -.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
SQ SEQUENCE 432 AA; 46385 MW; 2CE8D9A68E11945F CRC64;

Query Match          11.3%; Score 216.5; DB 5; Length 432;
Best Local Similarity 61.1%; Pred. No. 7.7e-09;
Matches 44; Conservative 3; Mismatches 12; Indels 13; Gaps 1;

Qy 26 HHHKKLQPCDGNPSPCVTCG-----KPKDELDYENDIEKKICKMEKC 72
   |::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 350 HIEQYLLKIQNSLSTWSPCVTCGNGIQVRIKPGSANKPKDELDYENDIEKKICKMEKC 409

Qy 73 SSVFNVNSNG 84
```

```
Db 410 SSVFNVNSSIG 421

RESULT 6
ID Q9U004 PRELIMINARY; PRT; 80 AA.
AC Q9U004;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BI;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269941; CAB64167.1; -.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9102 MW; A3283B70CEE50FDE CRC64;

Query Match          11.2%; Score 215.5; DB 5; Length 80;
Best Local Similarity 63.8%; Pred. No. 1.3e-09;
Matches 44; Conservative 1; Mismatches 11; Indels 13; Gaps 1;

Qy 26 HHHKKLQPCDGNPSPCVTCG-----KPKDELDYENDIEKKICKMEKC 72
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 12 HIEYLLKRIQNSLSTWSPCVTCGNGIQVRIKPGSANKPKDELDYENDIEKKICKMEKC 71

Qy 73 SSVFNVNS 81
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 72 SSVFNVNS 80

RESULT 7
ID Q9U0P1 PRELIMINARY; PRT; 80 AA.
AC Q9U0P1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BI896;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269978; CAB64197.1; -.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
```

Mon Feb 4 15:23:42 2002

```
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9102 MW; BFC6C970CEFOFA3E CRC64;

Query Match 11.2%; Score 215.5; DB 5; Length 80;
Best Local Similarity 63.8%; Pred. No. 1.3e-09;
Matches 44; Conservative 1; Mismatches 11; Indels 13; Gaps 1;

QY 26 HHKHKHKLKPGDGNPSPCVTCG-----KPKDELVDYNDIEKKICKMEKC 72
Db 12 HIEQYLKRIQNSLSTWSPCVTCGNGIQVRKPGSANKPKDQDLYNDIEKKICKMEKC 71
QY 73 SSVFNVVNSG 81
Db 72 SSVFNVVNSG 80

RESULT 8
ID Q9U934 PRELIMINARY; PRT; 115 AA.
AC Q9U934;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RC STRAIN=822A;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83154; AAA29567.1;
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 80 80
SQ SEQUENCE 115 AA; 12925 MW; A24CB33AEBE3F232 CRC64;

Query Match 11.1%; Score 213.5; DB 5; Length 115;
Best Local Similarity 59.7%; Pred. No. 2.9e-09;
Matches 43; Conservative 4; Mismatches 12; Indels 13; Gaps 1;

QY 26 HHKHKHKLKPGDGNPSPCVTCG-----KPKDELVDYNDIEKKICKMEKC 72
Db 33 HIEQYLKRIQNSLSTWSPCVTCGNGIQVRKPGSANKPKDQDLYNDIEKKICKMEKC 92
QY 73 SSVFNVVNSG 84
Db 93 SSVFNVVNSG 104

RESULT 9
ID Q25795 PRELIMINARY; PRT; 117 AA.
AC Q25795;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;

RN SEQUENCE FROM N.A.
RP SMART; SM00209; TSP1; 1.
RT "Geographically restricted heterogeneity of the Plasmodium falciparum circumsporozoite protein: relevance for vaccine development.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M77203; AAA29517.2;
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13110 MW; D96BE20944A7C726 CRC64;

Query Match 11.1%; Score 213.5; DB 5; Length 117;
Best Local Similarity 59.7%; Pred. No. 2.9e-09;
Matches 43; Conservative 4; Mismatches 12; Indels 13; Gaps 1;

QY 26 HHKHKHKLKPGDGNPSPCVTCG-----KPKDELVDYNDIEKKICKMEKC 72
Db 35 HIEQYLKRIQNSLSTWSPCVTCGNGIQVRKPGSANKPKDQDLYNDIEKKICKMEKC 94
QY 73 SSVFNVVNSG 84
Db 95 SSVFNVVNSG 106

RESULT 10
ID Q25829 PRELIMINARY; PRT; 416 AA.
AC Q25829;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RP STRAIN=MAD20;
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAD20;
RC la Cruz V.F.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MAD20;
RC Jongwutiwes S., Tanabe K., Kanbara H.;
RL Mol. Biochem. Parasitol. 0:0-0(0).
DR EMBL; M83172; AAA29550.1;
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
SQ SEQUENCE 416 AA; 44829 MW; D3EF560B2D368DE9 CRC64;

Query Match 11.1%; Score 213.5; DB 5; Length 416;
Best Local Similarity 59.7%; Pred. No. 1.3e-08;
```





DR PROSITE; PS50092; TSP1; 1.  
DR SMART; SM00209; TSP1; 1.  
FT NON\_TER 1  
SQ SEQUENCE 115 AA; 13018 MW; COA23F5805688237 CRC64;

Query Match 11.1%; Score 212.5; DB 5; Length 115;  
Best Local Similarity 59.7%; Pred. No. 3.4e-09;  
Matches 43; Conservative 4; Mismatches 12; Indels 13; Gaps 1;

QY 26 HHHKHKLKQPGDGNPWSVTCG-----KPKDELDYENDIEKKICKMEKC 72  
Db 33 HIEQYLKKIQYLSLSTWSPCSVTCGNGIQVRIKPGSADKPKDQLDYENDIEKKICKMEKC 92  
QY 73 SSVFNVVNSNG 84  
Db 93 SSVFNVVNSSIG 104

RESULT 14  
Q25839 PRELIMINARY; PRT; 115 AA.  
AC Q25839;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).  
GN CSP.  
OS Plasmodium falciparum (isolate K1 / Thailand).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K1;  
RX MEDLINE=84250215; PubMed=6204383;  
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
RA Hockmeyer W.T., Mayo W.L., Haynes J.D., Schneider I., Roberts D.,  
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;  
RT "Structure of the gene encoding the immunodominant surface antigen on  
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";  
RL Science 225:593-599(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K1;  
RA la Cruz V.F.;  
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K1;  
RA Jongwutiwes S., Tanabe K., Kanbara H.;  
RL Mol. Biochem. Parasitol. 0:0-0(0).  
DR EMBL; M83174; AAA29552.1; -;  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR003067; Crmsprzoite.  
DR Pfam; PF00090; tsp.1; 1.  
DR PRINTS; PR01303; CRCMSPRZOITE.  
DR PROSITE; PS50092; TSP1; 1.  
DR SMART; SM00209; TSP1; 1.  
FT NON\_TER 1  
SQ SEQUENCE 420 AA; 45318 MW; F7F70F1C4939DEA7 CRC64;

Query Match 11.1%; Score 212.5; DB 5; Length 420;  
Best Local Similarity 59.7%; Pred. No. 1.5e-08;  
Matches 43; Conservative 4; Mismatches 12; Indels 13; Gaps 1;

QY 26 HHHKHKLKQPGDGNPWSVTCG-----KPKDELDYENDIEKKICKMEKC 72  
Db 338 HIEQYLKKIQYLSLSTWSPCSVTCGNGIQVRIKPGSADKPKDQLDYENDIEKKICKMEKC 397  
QY 73 SSVFNVVNSNG 84  
Db 398 SSVFNVVNSSIG 409

Search completed: January 29, 2002, 11:12:07  
Job time: 763 sec

DR PROSITE; PS50092; TSP1; 1.  
DR SMART; SM00209; TSP1; 1.  
FT NON\_TER 1  
SQ SEQUENCE 115 AA; 13018 MW; COA23F5805688237 CRC64;

Query Match 11.1%; Score 212.5; DB 5; Length 115;  
Best Local Similarity 59.7%; Pred. No. 3.4e-09;  
Matches 43; Conservative 4; Mismatches 12; Indels 13; Gaps 1;

QY 26 HHHKHKLKQPGDGNPWSVTCG-----KPKDELDYENDIEKKICKMEKC 72  
Db 33 HIEQYLKKIQYLSLSTWSPCSVTCGNGIQVRIKPGSADKPKDQLDYENDIEKKICKMEKC 92  
QY 73 SSVFNVVNSNG 84  
Db 93 SSVFNVVNSSIG 104

RESULT 14  
Q25839 PRELIMINARY; PRT; 115 AA.  
AC Q25839;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).  
GN CSP.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=835C;  
RX MEDLINE=84250215; PubMed=6204383;  
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
RA Hockmeyer W.T., Mayo W.L., Haynes J.D., Schneider I., Roberts D.,  
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;  
RT "Structure of the gene encoding the immunodominant surface antigen on  
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";  
RL Science 225:593-599(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=835C;  
RA la Cruz V.F.;  
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=835C;  
RA Jongwutiwes S., Tanabe K., Kanbara H.;  
RL Mol. Biochem. Parasitol. 0:0-0(0).  
DR EMBL; M83162; AAA29575.1; -;  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR003067; Crmsprzoite.  
DR Pfam; PF00090; tsp.1; 1.  
DR PRINTS; PR01303; CRCMSPRZOITE.  
DR PROSITE; PS50092; TSP1; 1.  
DR SMART; SM00209; TSP1; 1.  
FT NON\_TER 1  
SQ SEQUENCE 115 AA; 12974 MW; 6BB538287260DA90 CRC64;

Query Match 11.1%; Score 212.5; DB 5; Length 115;  
Best Local Similarity 59.7%; Pred. No. 3.4e-09;  
Matches 43; Conservative 4; Mismatches 12; Indels 13; Gaps 1;

QY 26 HHHKHKLKQPGDGNPWSVTCG-----KPKDELDYENDIEKKICKMEKC 72  
Db 33 HIEQYLKKIQYLSLSTWSPCSVTCGNGIQVRIKPGSADKPKDQLDYENDIEKKICKMEKC 92  
QY 73 SSVFNVVNSNG 84  
Db 93 SSVFNVVNSSIG 104



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:07 ; Search time 285.36 Seconds  
(without alignments)  
8.201 Million cell updates/sec

Title: US-09-763-397A-3

Perfect score: 97  
Sequence: 1 KPLDKFGNIYDYHYEH 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_17:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	217	5	Q27336 plasmodium
2	96	99.0	217	5	P90588 plasmodium
3	48	49.5	246	2	Q92708 chlamydia p
4	48	49.5	1512	10	Q9M0S8
5	47	48.5	227	6	Q97909
6	46	47.4	207	5	Q02089
7	45	46.4	339	5	Q20045
8	45	46.4	779	10	Q9C6W4
9	44.5	45.9	793	2	O85356
10	44	45.4	336	5	Q18366
11	44	45.4	374	13	Q9YHX1
12	44	45.4	656	10	Q9ATV7
13	44	45.4	1115	3	Q9HGE0
14	44	45.4	1404	10	Q9T011
15	44	45.4	1517	10	Q9SEG3
16	44	45.4	1519	10	Q23273
17	43	44.3	505	8	Q9TIM6
18	43	44.3	600	2	Q9KQF7
19	43	44.3	1594	2	Q32608

Q9VJc6 drosophila  
Q83137 bean yellow  
Q65892 bean yellow  
Q9K5P1 bacillus ha  
Q95VU6 staphylococ  
Q26259 chironomus  
Q9JSY6 neisseria m  
Q9VTS4 drosophila  
Q9LWJ1 oryza sativ  
Q9P5V3 neurospora  
Q9ADV7 uvularia se  
Q96240 plasmodium  
Q04281 phaseolus v  
Q23814 calliphora  
O18320 calliphora  
Q9F0K1 staphylococ  
Q9X671 entamoeba h  
Q9SAL2 arabidopsis  
Q24847 entamoeba h  
Q9S5C4 unidentified  
Q9XDZ1 unidentified  
Q9XVV9 unidentified  
Q9ULZ8 caenorhabdi  
O32856 marinomonas  
Q9S093 borrelia bu  
Q91qt6 arabidopsis

43 44.3 4010 5 Q9VJc6  
42.5 43.8 1377 12 Q83137  
42.5 43.8 3056 12 Q65892  
42 43.3 169 2 Q9K5P1  
42 43.3 229 2 Q95VU6  
42 43.3 267 5 Q26259  
42 43.3 273 2 Q9JSY6  
42 43.3 406 5 Q9VTS4  
42 43.3 474 10 Q9LWJ1  
42 43.3 490 3 Q9P5V3  
42 43.3 658 8 Q9ADV7  
42 43.3 1182 5 Q96240  
41.5 42.8 385 10 Q04281  
41.5 42.8 668 5 Q23814  
41.5 42.8 759 5 O18320  
41 42.3 127 2 Q9F0K1  
41 42.3 273 5 Q9X671  
41 42.3 274 10 Q9SAL2  
41 42.3 294 5 Q24847  
41 42.3 321 2 Q9S5C4  
41 42.3 323 2 Q9XDZ1  
41 42.3 323 2 Q9XVV9  
41 42.3 349 5 Q9ULZ8  
41 42.3 422 2 O32856  
41 42.3 444 2 Q9S093  
41 42.3 483 10 Q91qt6

#### ALIGNMENTS

RESULT 1  
Q27336  
ID Q27336 PRELIMINARY; PRT; 217 AA.  
AC Q27336;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
DE PFG27/25.  
GN PFG 27-25.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91304524; PubMed=1852178;  
RA Alano P., Premawansa S., Bruce M.C., Carter R.;  
RT "A stage specific gene expressed at the onset of gametocytogenesis in Plasmodium falciparum.";  
RL Mol. Biochem. Parasitol. 46:81-88(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97008413; PubMed=8855557;  
RA Alano P., Silvestrini F., Roca L.;  
RT "Structure and polymorphism of the upstream region of the pfg27/25 gene, transcriptionally regulated in gametocytogenesis of Plasmodium falciparum.";  
RL Mol. Biochem. Parasitol. 79:207-218(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=FC1/HN;  
RA Li X.R., Yu X.B., Luo S.H., Shan Z.X., Fang J.M., Xu J.;  
RT "Sequence of Plasmodium falciparum encoding pfg27/25.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X84904; CAA59328.1; -;  
DR EMBL; AF179422; AAD55784.1; -;  
SQ SEQUENCE 217 AA; 26004 MW; B19C47E9B191F7A5 CRC64;

Query Match 100.0%; Score 97; DB 5; Length 217;  
Best Local Similarity 100.0%; Pred. No. 1.2e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDYHYEH 16  
 |||||  
 Db 10 KPLDKFGNIYDYHYEH 25

RESULT 2  
 ID P90588 PRELIMINARY; PRT; 217 AA.  
 AC P90588;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE GAMATE ANTIGEN (FRAGMENT).  
 OS Plasmodium reichenowi.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5854;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93295437; PubMed=8515780;  
 RA Lal A.A., Goldman I.F., Collins W.E., Kumar N.;  
 RT "Sequence of a 27-kilodalton gamete antigen of Plasmodium reichenowi  
 and comparison with Pfgr27 of Plasmodium falciparum.";  
 RL Mol. Biochem. Parasitol. 59:175-176(1993).  
 DR EMBL: L08799; AAB42050.1; -;  
 FT NON TER 217  
 SQ SEQUENCE 217 AA; 25943 MW; 24945C26C2252AC2 CRC64;

Query Match 99.0%; Score 96; DB 5; Length 217;  
 Best Local Similarity 93.8%; Pred. No. 1.8e-07;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDYHYEH 16  
 |||||  
 Db 10 KPLDKFGNIYDYHYEH 25

RESULT 3  
 ID Q92708 PRELIMINARY; PRT; 246 AA.  
 AC Q92708;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CT488 HYPOTHETICAL PROTEIN.  
 GN CPN0606 OR CPJ0606 OR CP0141.  
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=83558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=CWL029;  
 RC MEDLINE=99206606; PubMed=10192388;  
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,  
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";  
 RL Nat. Genet. 21:385-389(1999).  
 DR [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J138;  
 RX MEDLINE=20330349; PubMed=10871362;  
 RA Shira M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
 RA Shiba T., Ishii K., Hatatori M., Kuhara S., Nakazawa T.;  
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 from Japan and CWL029 from USA.";  
 RL Nucleic Acids Res. 28:2311-2314(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AR39;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Umayah L.A., Utterback T.,  
 RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,

Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J.,  
 McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;  
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia  
 pneumoniae Ar39."; 28:1397-1406(2000).  
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 DR EMBL: AE001645; AAD18745.1; -;  
 DR EMBL: AP002547; BAA98813.1; -;  
 DR EMBL: AE002175; AAF38023.1; -;  
 DR TIGR: CP0141;  
 DR InterPro: IPR000934; Ser\_thr\_phosphatase.  
 KW Complete proteome.  
 SQ SEQUENCE 246 AA; 27619 MW; F3990B645AA0083B CRC64;

Query Match 49.5%; Score 48; DB 2; Length 246;  
 Best Local Similarity 57.1%; Pred. No. 7.6;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDYHY 14  
 :|||  
 Db 217 REPIDFGNGIRGIHY 230

RESULT 4  
 ID Q9M0S8 PRELIMINARY; PRT; 1512 AA.  
 AC Q9M0S8;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE MODIFICATION METHYLASE (EC 2.1.1.73) (CYTOSINE-SPECIFIC  
 METHYLTRANSFERASE).  
 GN AT4G08990.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,  
 RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,  
 RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,  
 RA Kewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA CYTOSINE - S-  
 CC ADENOSYL-L-HOMOCYSTEINE + DNA 5-METHYLCYTOSINE.  
 CC -1- SIMILARITY: TO C-5 CYTOSINE-SPECIFIC DNA METHYLASE FAMILY.  
 DR EMBL: AL161513; CAB78023.1; -;  
 DR InterPro: IPR001025; BAH.  
 DR InterPro: IPR001525; C5\_DNA\_meth.  
 DR Pfam: PF01426; BAH; 2.  
 DR Pfam: PF00145; DNA\_methylase; 3.  
 DR PRINTS: PR00105; C5METTRFRASE.  
 DR SMART: SM00439; BAH; 2.  
 DR PROSITE: PS00094; C5\_MTASE\_1; 1.  
 DR PROSITE: PS00095; C5\_MTASE\_2; 1.  
 KW Methyltransferase; Transferase.  
 SQ SEQUENCE 1512 AA; 171154 MW; D6D9BCC70DFAC0AE CRC64;

Query Match 49.5%; Score 48; DB 10; Length 1512;  
 Best Local Similarity 53.3%; Pred. No. 56;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDYHYE 15  
 |||  
 Db 159 KPASKYKKIYDYFFE 173



Query Match	Score 45;	DB 5;	Length 339;
Best Local Similarity	54.5%;	Pred. No. 32;	
Matches	6;	Conservative	4; Mismatches
QY	2	PLDKFGNIYDY 12	
DBD	188	PIDPFGSLFDY 198	
RESULT	8		
Q9C6W4		PRELIMINARY;	PRT; 779 AA.
AC	Q9C6W4		
DT	01-JUN-2001	(TRENBLrel. 17, Created)	
DT	01-JUN-2001	(TRENBLrel. 17, Last sequence update)	
DT	01-JUN-2001	(TRENBLrel. 17, Last annotation update)	
DT	BETA-GALACTOSIDASE (EC 3.2.1.23) (LACTASE).		
GN	F27M3_5.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.		
OX	NCBI_TaxID=3702;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=CV. COLUMBIA;		
RC	MEDLINE=21016719; PubMed=11130712;		
RA	White O., Alonso J., Altfeld H., Araujo R., Bowman C.L., Brooks S.Y.,		
RA	White O., Alonso J., Altfeld H., Araujo R., Bowman C.L., Brooks S.Y.,		
RA	Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,		
RA	Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,		
RA	Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,		
RA	Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,		
RA	Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,		
RA	Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,		
RA	Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,		
RA	Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,		
RA	Millischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,		
RA	Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,		
RA	Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,		
RA	Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,		
RA	Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,		
RA	Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;		
RT	"Sequence and analysis of chromosome 1 of the plant Arabidopsis		
RT	thaliana."		
RL	Nature 408:816-820(2000).		
CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-		
CC	-1- GALACTOSE RESIDUES IN BETA-D-GALACTOSIDES.		
CC	EMBL; AC074360; AAG60136.1; -		
CC	InterPro; IPR000922; Gal-lectin.		
DR	InterPro; IPR001944; Glyco_hydro_35.		
DR	Pfam; PF02140; Gal-lectin; 1.		
DR	Pfam; PF01301; Glyco_hydro_35; 1.		
DR	PRINTS; PR00742; GLHYDLASE35.		
DR	PRODOM; PD005612; Gal-lectin; 1.		
DR	PROSITE; PS01182; GLYCOSYL_HYDROL_F35; 1.		
DR	PROSITE; PS02028; SUEL-LECTIN; 1.		
KW	Glycosidase; Hydrolase; Hypothetical protein.		
QY	SEQUENCE	779 AA; 86342 MW; 888813DF31889080 CRC64;	
Query Match	46.48;	Score 45;	DB 10; Length 779;
Best Local Similarity	53.3%;	Pred. No. 80;	
Matches	8;	Conservative	2; Mismatches
QY	2	PLDKFGNIYDYHVEH 16	
DB	312	PLDFEGNLNQPKYGH 326	



Query Match 45.4%; Score 44; DB 5; Length 336;  
Best Local Similarity 60.0%; Pred. No. 46;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 FGNVYHYH 15  
DB 189 FGSVYHYHQ 198  
||:| |||

## RESULT 11

Q9VHX1 PRELIMINARY; PRT; 374 AA.  
AC Q9VHX1;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE NEUROPEPTIDE Y/PEPTIDE YY RECEPTOR YB.  
GN NPVRB.  
OS Gadus morhua (Atlantic cod).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae;  
OC Gadus.  
OX NCBI\_TaxID=8049;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Arvidsson A.-K., Wraith A., Jonsson-Rylander A.-C., Larhammar D.;  
RT "Cloning of a neuropeptide Y/peptide YY receptor from the Atlantic  
RT cod: the Yb receptor.";  
RL Regul. Pept. 0:0-0(1998).  
DR EMBL; AF073925; AAD02833.1; -.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm.1.1.  
DR PRINTS; PR00237; GPCRHHODPSN.  
DR PROSITE; PS0237; G-PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
DR PROSITE; PS0262; G-PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor; Neuropeptide.  
SQ SEQUENCE 374 AA; 42360 MW; D2A9C9516C4998E2 CRC64;

Query Match 45.4%; Score 44; DB 13; Length 374;  
Best Local Similarity 42.9%; Pred. No. 52;  
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 PLDKFGNIYDYHYE 15  
DB 278 PLNIENVFDMHHE 291  
||:| ||:| |||

## RESULT 12

Q9ATV7 PRELIMINARY; PRT; 656 AA.  
AC Q9ATV7;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE ARABINOXYLAN ARABINOFURANOHYDROLASE ISOENZYME AXAH-II.  
OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
OC Triticeae; Hordeum.  
OX NCBI\_TaxID=4513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lee R.C., Burton R.A., Hrmova M., Fincher G.B.;  
RT "Barley arabinoxylan arabinofuranohydrolases.";  
RL Biochem. J. 0:0-0(2001).  
DR EMBL; AF320325; AAK21880.1; -.  
KW Hydrolase.  
SQ SEQUENCE 656 AA; 71999 MW; 4D04531E5948415A CRC64;

Query Match 45.4%; Score 44; DB 10; Length 656;  
Best Local Similarity 53.8%; Pred. No. 96;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KPLDKFGNIYDYH 13  
DB 414 KPLDHPADLYDFH 426  
|||| :|||

## RESULT 13

Q9HGE0 PRELIMINARY; PRT; 1115 AA.  
AC Q9HGE0;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE FUM6P.  
GN FUM6.  
OS Gibberella moniliformis.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Hypocreales; Nectriaceae; Gibberella.  
OX NCBI\_TaxID=117187;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M-3125;  
RA Proctor R.H., Seo J.-A., Plattner R.D.;  
RT "Characterization of four clustered and coregulated genes associated  
RT with fumonisin biosynthesis in Fusarium verticillioides.";  
RL Submitted (Oct-2000) to the EMBL/Genbank/DDSI databases.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; AF155773; AAG27132.1; -.  
DR InterPro; IPR001128; Cyt.P450.  
DR InterPro; IPR003097; FAD\_binding.  
DR InterPro; IPR001094; Flavdxn-like.  
DR InterPro; IPR001709; Flavpyrid\_cyt\_redtse.  
DR InterPro; IPR001433; Oxidored\_FAD.  
DR Pfam; PF00667; FAD\_binding; 1.  
DR Pfam; PF00175; oxidored\_fad; 1.  
DR Pfam; PF00067; p450; 1.  
DR PRINTS; PR00369; FLAVODOXIN.  
DR PRINTS; PR00371; FPNCR.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS0086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 1115 AA; 123276 MW; 5C6D2B947AE86C25 CRC64;

Query Match 45.4%; Score 44; DB 3; Length 1115;  
Best Local Similarity 37.5%; Pred. No. 1.7e+02;  
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 KPLDKFGNIYDYHYEH 16  
DB 26 KPLPLGLNLFDFDN 41  
||| ||::| :

## RESULT 14

Q9T011 PRELIMINARY; PRT; 1404 AA.  
AC Q9T011;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE MODIFICATION METHYLASE (EC 2.1.1.73) (CYTOSINE-SPECIFIC  
DE METHYLTRANSFERASE).  
GN T6G15.160 OR AT4G13610.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.

Mon Feb 4 15:23:48 2002

DR PROSITE: PS00018; EF\_HAND; UNKNOWN\_1.  
 KW Methyltransferase; Transferase.  
 SQ SEQUENCE 1517 AA; 171324 MW; AF08810C3AFDB54A CRC64;  
 Query Match 45.4%; Score 44; DB 10; Length 1517;  
 Best Local Similarity 46.7%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 KPLDKFGNIYDYHYE 15  
 Db 159 KPSKKYKLYDYFFE 173  
 Search completed: January 29, 2002, 11:12:09  
 Job time: 765 sec

RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,  
 RA Mayer K.F.X., Schueller C.;  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA CYTOSINE = S-  
 CC ADENOSYL-L-HOMOCYSTEINE + DNA 5-METHYLCYTOSINE.  
 CC -1- SIMILARITY: TO C-5 CYTOSINE-SPECIFIC DNA METHYLASE FAMILY.  
 DR EMBL; AL049656; CAB41119.1; -.  
 DR EMBL; AL161537; CAB78403.1; -.  
 DR InterPro; IPR001025; BAH.  
 DR InterPro; IPR001525; C5\_DNA\_meth.  
 DR Pfam; PF01426; BAH; 2.  
 DR Pfam; PF00145; DNA\_methylase; 4.  
 DR PRINTS; PR00105; C5METHTRFRASE.  
 DR SMART; SM00439; BAH; 2.  
 DR PROSITE; PS00094; C5\_MTASE\_1; 1.  
 DR PROSITE; PS00095; C5\_MTASE\_2; 1.  
 DR Methyltransferase; Transferase.  
 KW SEQUENCE 1404 AA; 160211 MW; 2FC7B7AA8E5DCA8 CRC64;  
 SQ

Query Match 45.4%; Score 44; DB 10; Length 1404;  
 Best Local Similarity 46.7%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 KPLDKFGNIYDYHYE 15  
 Db 102 KPSKKYKLYDYFFE 116

RESULT 15  
 Q9SEG3 PRELIMINARY; PRT; 1517 AA.  
 AC Q9SEG3;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)  
 DE MODIFICATION METHYLASE (EC 2.1.1.73) (CYTOSINE-SPECIFIC  
 DE METHYLTRANSFERASE).  
 GN METII.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=CV. COLUMBIA;  
 RC MEDLINE=20044094; PubMed=10579493;  
 RX Genger R.K., Kovac K.A., Dennis E.S., Peacock W.J., Finnegan E.J.;  
 RA "Multiple DNA methyltransferase genes in Arabidopsis thaliana.";  
 RL Plant Mol. Biol. 41:269-278(1999).  
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA CYTOSINE = S-  
 CC ADENOSYL-L-HOMOCYSTEINE + DNA 5-METHYLCYTOSINE.  
 CC -1- SIMILARITY: TO C-5 CYTOSINE-SPECIFIC DNA METHYLASE FAMILY.  
 DR EMBL; AF138283; AAF14882.1; -.  
 DR HSP; P20589; 1DCT.  
 DR InterPro; IPR001025; BAH.  
 DR InterPro; IPR001525; C5\_DNA\_meth.  
 DR InterPro; IPR002048; EF-hand.  
 DR Pfam; PF01426; BAH; 2.  
 DR Pfam; PF00145; DNA\_methylase; 3.  
 DR PRINTS; PR00105; C5METHTRFRASE.  
 DR SMART; SM00439; BAH; 2.  
 DR PROSITE; PS00094; C5\_MTASE\_1; 1.  
 DR PROSITE; PS00095; C5\_MTASE\_2; 1.  
 DR



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:33 ; Search time 144.96 Seconds  
(without alignments)  
8.408 Million cell updates/sec

Title: US-09-763-397A-3

Perfect score: 97

Sequence: 1 KPLDKFGNIYDYHYEH 16

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_68:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	217	2 S27829	gametocytogenesis
2	48	49.5	246	2 H72057	conserved hypothet
3	48	49.5	246	2 C86566	CT488 hypothetical
4	48	49.5	1512	2 G85090	hypothetical prote
5	46	47.4	207	2 T34041	hypothetical prote
6	45	46.4	339	2 T16273	hypothetical prote
7	45	46.4	392	2 C72427	pyruvate synthase
8	44.5	45.9	316	2 G65021	transaldolase (EC
9	44.5	45.9	316	2 A85889	transaldolase A [l
10	44.5	45.9	507	2 S23422	catalase (EC 1.1.1.
11	44	45.4	336	2 T19665	hypothetical prote
12	44	45.4	591	1 CBBY2	L-lactate dehydrog
13	44	45.4	867	2 JH0225	L96 protein - Tipu
14	44	45.4	1072	2 A38457	integrin alpha-5 c
15	44	45.4	1404	2 T06663	DNA (cytosine-5)-
16	44	45.4	1519	2 G71402	DNA (cytosine-5)-
17	44	45.4	1534	2 S59604	DNA (cytosine-5)-
18	43	44.3	358	1 A48952	triacylglycerol li
19	43	44.3	540	2 A70358	topoisomerase I -
20	43	44.3	600	2 C92126	conserved hypothet
21	43	44.3	1594	2 T33072	hemolysin A - Edw
22	42	43.3	169	2 G84155	hypothetical prote
23	42	43.3	273	2 G81778	probable integral
24	42	43.3	395	2 T15302	hypothetical prote
25	42	43.3	490	2 T49616	hypothetical prote
26	42	43.3	1182	2 G71607	probable integral
27	42	43.3	1683	2 S56811	probable membrane
28	41.5	42.8	385	2 T11848	gibberellin 20-oxi
29	41.5	42.8	687	2 JQ1044	arylphorin precurs

30	41	42.3	273	2 B37281	replication factor
31	41	42.3	274	2 D96840	hypothetical prote
32	41	42.3	294	1 S58472	lysine-rich surfac
33	41	42.3	503	2 A40843	cytochrome P450 3
34	41	42.3	575	2 S39484	DNA-binding protei
35	41	42.3	603	2 F96797	probable GT-like t
36	41	42.3	720	2 F71660	dyein heavy chain
37	41	42.3	1114	2 T30299	hypothetical prote
38	40.5	41.8	798	2 T38490	hypothetical prote
39	40	41.2	87	2 T07456	deoxynucleotide mo
40	40	41.2	241	1 K1BPD4	bioc protein homol
41	40	41.2	251	1 B64456	probable transcrit
42	40	41.2	332	2 T41149	probable transcrit
43	40	41.2	358	2 S15908	site-specific DNA-
44	40	41.2	364	2 C70322	3-isopropylmalate
45	40	41.2	398	2 T27670	hypothetical prote

ALIGNMENTS

RESULT 1  
S27829  
gametocytogenesis onset-specific protein - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jun-2000  
C:Accession: S27829  
R:Alano, P.; Premawansa, S.; Bruce, M.; Carter, R.  
submitted to the EMBL Data Library, April 1991  
A:Description: A stage specific gene expressed at the onset of gametocytogenesis in P  
A:Reference number: S27829  
A:Accession: S27829  
A:Molecule type: mRNA  
A:Residues: 1-217 <ALA>  
A:Cross-references: GB:X84904; EMBL:M38286; NID:gl340125; PID:el39902; PID:gl340126

Query Match 100.0%; Score 97; DB 2; Length 217;  
Best Local Similarity 100.0%; Pred. No. 4.6e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPLDKFGNIYDYHYEH 16  
|||||  
DB 10 KPLDKFGNIYDYHYEH 25

RESULT 2  
H72057  
conserved hypothetical protein CP0141 [imported] - Chlamydophila pneumoniae (strains  
N:Alternate names: ct488 hypothetical protein  
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
C:Accession: H72057; C81609  
R:Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606  
A:Accession: H72057  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-246 <ARN>  
A:Cross-references: GB:AE001645; GB:AE001363; NID:g4376896; PIDN:AAI8745.1; PID:g437  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
A:Reference number: A81500; MUID:20150255  
A:Accession: C81609  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-246 <REA>  
A:Cross-references: GB:AE002175; GB:AE002161; NID:g7189069; PIDN:AAF38023.1; PID:g718

N:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchii, K.; Shiba, I.; R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchii, K.; Shiba, I.;  
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
 A:Reference number: A86491; MUID:20330349  
 A:Accession: C86566  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-245 <STO>  
 A:Cross-references: GB:BA000008; NID:g8578978; PIDN:BA098813.1; GSPDB:GN00142  
 A:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: CP10606

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #seq\_change 10-Sep-1999  
 C:Accession: F16273  
 R:Connell, M.  
 submitted to the EMBL Data Library, June 1995  
 A:Description: The sequence of *C. elegans* cosmid F35D2.  
 A:Reference number: Z18488  
 A:Accession: F16273  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-339 <CON>  
 A:Cross-references: EMBL:U08741; NID:g861290; PID:g861293; PIDN:AAA68327.1  
 A:Experimental source: strain Bristol N2  
 C:Genetics:  
 A:Gene: CESP:F35D2.1  
 A:Recons: 76/1; 102/1; 136/3; 174/1; 240/2; 268/1; 298/1

N; Alternate names: pyruvate-ferredoxin 2-oxido-reductase alpha subunit  
C: Species: *Thermotoga maritima*  
C: Date: 11-Jun-1999 #sequence\_revision 18-Jun-1999 #text\_change 21-Jul-2000  
C: Accession: C74227 #A54346  
R: Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.;  
C.M.  
Nature 399, 323-329, 1999  
A: title: Evidence for lateral gene transfer between Archaea and Bacteria

A:Reference number: A72200; MUID:99287316

A:Accession: C72427

A:Molecule type: DNA

A:Residues: 'MR' 1-392 <NEL>

A:Cross-references: GB:AE001690; GB:AE000512; NID:g4980496; PIDN:AAD35111.1; PID:g498050

A:Experimental source: strain MSB8

A:Note: an incorrect initiation codon was used

R:Blamey, J.M.; Adams, M.W.

Biochemistry 33, 1000-1007, 1994

A:Title: Characterization of an ancestral type of pyruvate ferredoxin oxidoreductase from

A:Reference number: A54346; MUID:94137707

A:Accession: A54346

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-43 <BLA>

A:Experimental source: strain DSM 3109

A:Note: sequence extracted from NCBI backbone (NCBIP:143495)

C:Genetics:

A:Gene: TM0017

C:Superfamily: pyruvate synthase alpha chain; 2-oxoacid ferredoxin oxidoreductase homolog

C:Keywords: coenzyme A; oxidoreductase

Query Match 46.4%; Score 45; DB 2; Length 392;

Best Local Similarity 53.3%; Pred. No. 16;

Matches 8; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

Qy 2 PLDFGFGNIYDYHYEH 16

||| : ||| : ||| : |||

Db 214 PLD-----LYDYFEH 224

RESULT 8

G65021

transaldolase (EC 2.2.1.2) b2464 - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 12-Nov-1999

C:Accession: G65021

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: G65021

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-316 <BLAT>

A:Cross-references: GB:AE000333; GB:U00096; NID:gl788805; PIDN:AAC75517.1; PID:gl788807;

A:Experimental source: strain K-12, substrain MG1655

C:Function:

A:Description: catalyzes the reversible transfer of a three-carbon ketol unit from sedoh

C:Superfamily: human transaldolase

C:Keywords: transferase

F:131/Active site: Lys #status predicted

Query Match 45.9%; Score 44.5; DB 2; Length 316;

Best Local Similarity 33.3%; Pred. No. 15;

Matches 9; Conservative 4; Mismatches 3; Indels 11; Gaps 1;

Qy 1 KPLDKF-----GNIYDYHYEH 16

||| : ||| : ||| : |||

Db 189 KPMDPYVVEEDPGVKSVRNIVDYKQH 215

RESULT 9

A85889

transaldolase A [imported] - Escherichia coli (strain O157:H7)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001

C:Accession: A85889

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: A85889

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-316 <STO>

A:Cross-references: GB:AE005174; NID:gl2516839; PIDN:AAG57573.1; GSPDB:GN00145; UMG:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: tala

C:Superfamily: human transaldolase

Query Match 45.9%; Score 44.5; DB 2; Length 316;

Best Local Similarity 33.3%; Pred. No. 15;

Matches 9; Conservative 4; Mismatches 3; Indels 11; Gaps 1;

Qy 1 KPLDKF-----GNIYDYHYEH 16

||| : ||| : ||| : |||

Db 189 KPMDPYVVEEDPGVKSVRNIVDYKQH 215

RESULT 10

S23422

catalase (EC 1.11.1.6) - yeast (Pichia angusta)

C:Species: Pichia angusta

C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 20-Apr-2000

C:Accession: S23422

R:Didion, T.; Roggenkamp, R.

FEBS Lett. 303, 113-116, 1992

A:Title: Targeting signal of the peroxisomal catalase in the methylotrophic yeast Han

A:Reference number: S23422; MUID:92299073

A:Accession: S23422

A:Molecule type: DNA

A:Residues: 1-507 <DID>

A:Cross-references: EMBL:X56501; NID:g2775; PIDN:CAA39856.1; PID:g2776

C:Superfamily: catalase

C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase

F:65,104,138/Active site: His, Ser, Asn #status predicted

F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 45.9%; Score 44.5; DB 2; Length 507;

Best Local Similarity 47.1%; Pred. No. 26;

Matches 8; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

Qy 1 KPLDKF-GNIYDYHYEH 16

|| : || : || : ||

Db 415 KPDEKVTGEVVPYRWEH 431

RESULT 11

T19665

hypothetical protein C33A12.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000

C:Accession: T19665

R:Wilkinson, J.

submitted to the EMBL Data Library, January 1996

A:Reference number: Z19159

A:Accession: T19665

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-336 <WIL>

A:Cross-references: EMBL:Z68493; PIDN:CAA92795.1; GSPDB:GN00022; CESP:C33A12.11

A:Experimental source: clone C33A12

C:Genetics:

A:Gene: CESP:C33A12.11

A:Map position: 4

A:Introns: 149/3; 202/3; 270/3

C:Superfamily: Caenorhabditis elegans hypothetical protein C33A12.9b

Query Match 45.4%; Score 44; DB 2; Length 336;  
Best Local Similarity 60.0%; Pred. No. 19;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 FGNIDYHYE 15  
||:| |||  
Db 189 FGSVYVHYQ 198  
||:| |||

RESULT 12  
CBBY2  
L-lactate dehydrogenase (cytochrome) (EC 1.1.2.3) precursor - yeast (Saccharomyces cerevisiae)  
N:Alternate names: cytochrome b2; flavocytochrome B2; protein YML054  
C:Species: Saccharomyces cerevisiae  
C:Date: 24-Apr-1984 #sequence\_revision 31-Mar-1993 #text\_change 21-Jul-2000  
C:Accession: A24583; S49806; A91136; A23095; A90671; A00175  
R:Guilard, B.  
EMBO J. 4, 3265-3272, 1985  
A:Title: Structure, expression and regulation of a nuclear gene encoding a mitochondrial  
A:Reference number: A24583; MUID:86135959  
A:Accession: A24583  
A:Molecule type: DNA  
A:Residues: 1-591 <GUI1>  
A:Cross-references: EMBL:X03215; NID:g3632; PIDN:CAA26959.1; PID:g3633  
R:Devlin, K.; Churcher, C.  
submitted to the EMBL Data Library, November 1994  
A:Reference number: S49800  
A:Accession: S49806  
A:Molecule type: DNA  
A:Residues: 1-591 <DEV>  
A:Cross-references: EMBL:Z46729; NID:g577134; PIDN:CAA86721.1; PID:g577142; GSPDB:GN0001  
R:Ghrir, R.; Becam, A.M.; Lederer, F.  
Eur. J. Biochem. 139, 59-74, 1984  
A:Title: Primary structure of flavocytochrome b2 from baker's yeast.  
A:Reference number: A91136; MUID:84132029  
A:Accession: A91136  
A:Molecule type: protein  
A:Residues: 80,82-164,'E',166-394 <GHR>  
R:Guilard, B.; Lederer, F.; Jacq, C.  
Nature 255, 422-423, 1975  
A:Title: More similarity between bakers' yeast L-(-)-lactate dehydrogenase and liver mld  
A:Reference number: A93173; MUID:75156546  
A:Accession: A93173  
A:Molecule type: protein  
A:Residues: 'N',82-94 <GUI2>  
R:Lederer, F.; Cortial, S.; Becam, A.M.; Haumont, P.Y.; Perez, L.  
Eur. J. Biochem. 152, 419-428, 1985  
A:Title: Complete amino acid sequence of flavocytochrome b2 from baker's yeast.  
A:Reference number: A91154; MUID:86030284  
A:Accession: A23095  
A:Molecule type: protein  
A:Residues: 395-465,'Q',467-513,'E',515-591 <LED>  
R:Guilard, B.; Lederer, F.  
Biochimie 58, 305-316, 1976  
A:Title: Complete amino acid sequence of the heme-binding core in bakers' yeast cytochr  
A:Reference number: A90671; MUID:76206228  
A:Accession: A90671  
A:Molecule type: protein  
A:Residues: 88-121,'D',123-164,'E',166-183 <GUI3>  
R:Xia, Z.; Mathews, F.S.  
J. Mol. Biol. 212, 837-863, 1990  
A:Title: Molecular structure of flavocytochrome b-2 at 2.4 angstrom resolution.  
A:Reference number: A44532; MUID:90230315  
A:Contents: annotation; X-ray crystallography, 2.4 angstroms  
C:Genetics:  
A:Gene: SGD:CYB2; MIPS:YML054c  
A:Cross-references: SGD:S0004518; MIPS:YML054c  
A:Map position: 13L  
A:Genome: nuclear  
C:Superfamily: lactate dehydrogenase (cytochrome); (S)-2-hydroxy-acid oxidase homology;  
C:Keywords: chromoprotein; electron transfer; flavoprotein; FMN; heme; homotetramer; iron  
F:1-80/Domain: transit peptide (mitochondrion) #status predicted <TNP>

F:81-591/Product: L-lactate dehydrogenase (cytochrome) #status predicted <MAT>  
F:88-159/Domain: cytochrome b5 core homology <CB5>  
F:200-504/Domain: (S)-2-hydroxy-acid oxidase homology <CB5>  
F:123-146/Binding site: heme iron (His) (axial ligands) #status experimental  
F:429,459,493,513/Binding site: FMN (Lys, Asp, Arg, Arg) #status experimental  
F:453/Active site: His #status experimental

Query Match 45.4%; Score 44; DB 1; Length 591;  
Best Local Similarity 53.8%; Pred. No. 36;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 PLDKFGNIYDYHY 14  
||| |::| |  
Db 198 PLDNIINLYDFEY 210  
||| |::| |

## RESULT 13

JH0225  
L96 protein - Tipula iridescent virus  
C:Species: Tipula iridescent virus  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: JH0225  
R:Home, W.A.; Tajbakhsh, S.; Seligy, V.L.  
Gene 94, 243-248, 1990  
A:Title: Molecular cloning and characterization of a late Tipula iridescent virus gen  
A:Reference number: JH0225; MUID:91078646  
A:Accession: JH0225  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-867 <HOM>  
A:Cross-references: GB:M34853; NID:g335215; PIDN:AAA47919.1; PID:g335216

Query Match 45.4%; Score 44; DB 2; Length 867;  
Best Local Similarity 64.3%; Pred. No. 56;  
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KPLDKFGNIYDYHY 14  
||| |::| |  
Db 726 KPLPLGLGIDDDYHY 739  
||| |::| |

## RESULT 14

A38457  
Integrin alpha-6 chain precursor - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 29-Sep-1999  
C:Accession: A38457; S16429  
R:de Curtis, I.; Quaranta, V.; Tamura, R.N.; Reichardt, L.F.  
J. Cell Biol. 113, 405-416, 1991  
A:Title: Laminin receptors in the retina: sequence analysis of the chick integrin alp  
A:Reference number: A38457; MUID:91185416  
A:Accession: A38457  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1072 <DEC>  
A:Cross-references: GB:X56559; NID:g63541; PIDN:CAA39909.1; PID:g63542  
C:Superfamily: integrin alpha-2b chain  
C:Keywords: cell adhesion; cytoskeleton; phosphoprotein; transmembrane protein

Query Match 45.4%; Score 44; DB 2; Length 1072;  
Best Local Similarity 58.3%; Pred. No. 72;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 PLDKFGNIYDYHY 13  
||| |::| |  
Db 394 PYDGFQKVIYVH 405  
||| |::| |

## RESULT 15

T06663



DNA (cytosine-5-)-methyltransferase (EC 2.1.1.37) T6G15.160 - Arabidopsis thaliana  
N:Alternate names: protein T6G15.160  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 05-May-2000  
C:Accession: T06663  
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Meyer, K.F.X  
submitted to the Protein Sequence Database, April 1999  
A:Reference number: 215791  
A:Accession: T06663  
A:Molecule type: DNA  
A:Residues: 1-1404 <BEV>  
A:Cross-references: EMBL:AL049656; GSPDB:GN00062; ATSP:T6G15.160  
A:Experimental source: cultivar Columbia; BAC clone T6G15  
C:Genetics:  
A:Gene: ATSP:T6G15.160  
A:Map position: 4  
A:Introns: 57/1; 348/3; 371/2; 872/3; 927/3; 989/1; 1026/2; 1086/3; 1152/3; 1245/3; 1286/3  
C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 45.4%; Score 44; DB 2; Length 1404;  
Best Local Similarity 46.7%; Pred. No. 98;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDYHYE 15  
||| :||| :!  
Db 102 KPSKKYKKLYDYFFE 116

Search completed: January 29, 2002, 10:26:34  
Job time: 645 sec

**THIS PAGE BLANK (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:40 ; Search time 310.82 Seconds

(without alignments)  
3.813 Million cell updates/sec

Title: US-09-763-397A-3

Perfect score: 97

Sequence: 1 KPLDFGNIYDYHYEH 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	100.0	16	21	AAV70279
2	97	100.0	350	21	AAV70278
3	62	63.9	20	17	AAV96217
4	52	53.6	202	22	AAV81899
5	52	53.6	312	22	AAV82008
6	48	49.5	262	20	AAV35228
7	44	45.4	89	21	AAV71539
8	44	45.4	303	21	AAV71534
9	43	44.3	319	10	AAV91951
10	43	44.3	319	17	AAV09624
11	43	44.3	319	17	AAV09625
					Plasmodium falciparum
					Recombinant vaccinia
					Continuous epitope
					S. epidermidis ope
					S. epidermidis ope
					Amino acid sequenc
					Soybean POP2/CAFL
					Soybean POP2/CAFL
					Polypeptide with 1
					Pseudomonas glumae

12	43	44.3	319	17	AAV88018	Mature Pseudomonas
13	43	44.3	319	17	AAV88010	Mature Pseudomonas
14	43	44.3	319	17	AAV88011	Mature Pseudomonas
15	43	44.3	319	17	AAV88012	Mature Pseudomonas
16	43	44.3	319	17	AAV88013	Mature Pseudomonas
17	43	44.3	319	17	AAV88014	Mature Pseudomonas
18	43	44.3	319	17	AAV88015	Mature Pseudomonas
19	43	44.3	319	17	AAV88016	Mature Pseudomonas
20	43	44.3	319	17	AAV88017	Mature Pseudomonas
21	43	44.3	329	22	AAV47427	EPM. Lactococcus
22	43	44.3	339	15	AAV2656	E.coli ompA signal
23	43	44.3	358	12	AAV10614	Pseudomonas glumae
24	43	44.3	358	13	AAV20177	P.glumae PGI lipas
25	43	44.3	360	12	AAV10438	Pseudomonas glumae
26	43	44.3	361	12	AAV10437	Pseudomonas glumae
27	42	43.3	123	18	AAV28080	Staphylococcus aur
28	42	43.3	181	21	AAV32984	Pinus radiata tran
29	42	43.3	273	21	AAV74895	Neisseria meningit
30	42	43.3	1182	21	AAV18288	Plasmodium falcipa
31	41	42.3	271	22	AAV84859	Protein isolated f
32	41	42.3	275	22	AAV84858	Protein isolated f
33	41	42.3	550	21	AAV07900	Arabidopsis thalia
34	41	42.3	575	21	AAV07899	Arabidopsis thalia
35	41	42.3	597	21	AAV07898	Arabidopsis thalia
36	41	42.3	674	22	AAV00416	Tomato hydrolase,
37	40	41.2	107	21	AAV32819	Eucaalyptus grandis
38	40	41.2	375	22	AAV41751	Human polypeptide
39	40	41.2	375	22	AAV41752	Human polypeptide
40	40	41.2	445	21	AAV55790	Superheat-resiscan
41	40	41.2	604	21	AAV16602	Bacteriophage 192
42	40	41.2	638	22	AAV95680	Human protein sequ
43	40	41.2	662	22	AAV39966	Human polypeptide
44	40	41.2	683	22	AAV92931	C glutamicum prote
45	40	41.2	688	22	AAV39965	Human polypeptide

ALIGNMENTS

RESULT 1

AAV70279  
ID AAV70279 standard; peptide; 16 AA.

AC AAV70279;

DT 06-JUN-2000 (first entry)

DE Plasmodium falciparum Pfg27 antigenic epitope, P591.

KW Recombinant protein; CDC/NIAI/VAC-1; multivalent; malaria; vaccine;  
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;  
KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;  
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-2;  
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;  
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;  
KW Pfg27; antiparasitic; prevention; anti-CDC/NIAI/VAC-1 antibody.  
OS Plasmodium falciparum.

PN WO200011179-A1.

PD 02-MAR-2000.

PF 19-AUG-1999; 99WO-US18869.

PR 21-AUG-1998; 98US-0097703.

PA (NAIM-) NAT INST IMMUNOLOGY.  
(USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Lal AA, Shi YP, Hasnain SE;

DR WPI; 2000-237654/20.

Mon Feb 4 15:23:47 2002

us-09-763-397a-3.rag

```

XX Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT plasmodium falciparum life cycle
XX
XX Claim 2; Page 17; 52pp; English.
PS
XX The present sequence is the antigenic epitope p591, derived from
CC gamete specific antigen, Pf27 of the sexual stage of Plasmodium
CC falciparum. It is used in the construction of recombinant protein
CC CDC/NIH/VAC-1, which is a multivalent, multistage malarial vaccine.
CC The recombinant protein comprises, melittin signal peptide, (His)6 tag,
CC T-cell epitope from tetanus toxoid and 21 antigenic epitopes from
CC circumsporozoite protein (CSP), sporozoite surface protein-2 (SSP-2),
CC liver stage antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1),
CC MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding antigen-175
CC (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete specific
CC antigen, Pf27. These epitopes were obtained at different stages of the
CC life cycle of P. falciparum. CDC/NIH/VAC-1 vaccine has antiparasitic
CC activity and can be used for treatment and prevention of malarial
CC infections. Anti-CDC/NIH/VAC-1 antibodies can be used for detecting
CC P. falciparum in biological samples.
XX
XX Sequence 16 AA:
SQ
Query Match 100.0%; Score 97; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPLDKFGNIYDYHYEH 16
DB 1 kpldkfgniydyhyeh 16
RESULT 2
AAAY70278
ID AAAY70278 standard; Protein: 350 AA.
XX
XX AC AAAY70278;
XX
XX DT 06-JUN-2000 (first entry)
XX
XX DE Recombinant vaccine CDC/NIH/VAC-1.
XX
XX KW Recombinant protein; CDC/NIH/VAC-1; multivalent; malaria; vaccine;
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KW Pf27; antiparasitic; prevention; anti-CDC/NIH/VAC-1 antibody;
KW honey bee.
XX
XX OS Chimeric - Apis sp.
OS Chimeric - Clostridium tetani.
OS Chimeric - Plasmodium falciparum.
XX
XX FH Key Location/Qualifiers
XX Peptide 1..22
FT /label= Melittin_signal_peptide
FT /note= "Derived from Honey bee"
FT Protein 23..350
FT /label= Mature_CDC/NIH/VAC-1
FT /note= "Recombinant multivalent malarial vaccine"
XX
XX WO200011179-A1.
XX
XX PN 02-MAR-2000.
XX
XX PD 19-AUG-1999; 99WO-US18869.
XX
XX PF 21-AUG-1998; 98US-0097703.
XX

```

```

XX (NAIM-) NAT INST IMMUNOLOGY.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Lal AA, Shi YP, Hasnain SE;
PI
XX WPI; 2000-237654/20.
XX N-PSDB; AA251336.
DR
XX Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT plasmodium falciparum life cycle
XX
XX Claim 3; Page 43-44; 52pp; English.
PS
XX The present sequence is that of recombinant protein CDC/NIH/VAC-1,
CC which is a multivalent, multistage malarial vaccine. The recombinant
CC protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope
CC from tetanus toxoid and 21 antigenic epitopes from circumsporozoite
CC protein (CSP), sporozoite surface protein-1 (MSP-1), MSP-2, apical
CC antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical
CC membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),
CC rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Pf27.
CC These epitopes were obtained at different stages of the life cycle of
CC Plasmodium falciparum. CDC/NIH/VAC-1 vaccine has antiparasitic
CC activity and can be used for treatment and prevention of malarial
CC infections. Anti-CDC/NIH/VAC-1 antibodies can be used for detecting
CC P. falciparum in biological samples.
XX
XX Sequence 350 AA:
SQ
Query Match 100.0%; Score 97; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPLDKFGNIYDYHYEH 16
DB 283 kpldkfgniydyhyeh 298
RESULT 3
AAR96217
ID AAR96217 standard; peptide; 20 AA.
XX
XX AC AAR96217;
XX
XX DT 15-AUG-1996 (first entry)
XX
XX DE Continuous epitope of Plasmodium falciparum in the gametocyte stage.
XX Epitope: monoclonal antibody; malaria; vaccine; immunisation;
KW transmission; sexual stage.
KW
XX Plasmodium falciparum.
OS
XX US5502168-A.
XX
XX PD 26-MAR-1996.
XX
XX PF 24-OCT-1991; 91US-0779494.
XX
XX PR 24-OCT-1991; 91US-0779494.
XX
XX PR 14-SEP-1993; 93US-0120225.
XX
XX PA (UWJO ) UNIV JOHNS HOPKINS.
XX
XX FI Kumar N;
XX
XX WPI; 1996-179336/18.
XX
XX Monoclonal antibodies useful in vaccines against malaria
PT transmission - recognise continuous epitope in the sexual stages of
PT

```

CC polynucleotide sequences from the present invention. AAH55091 to  
CC AAH55098 represent oligonucleotide sequences and primers which are used  
CC in the exemplification of the present invention.  
CC N.B. The present invention specifically claims all the polynucleotide  
CC sequences given in the sequence listing of the present specification,  
CC however the sequence listing only goes up to SEQ ID NO:4454 so even  
CC though sequences are given in the disclosure for SEQ ID NO:4472, no  
CC no sequences are present for SEQ ID NO:4455 to 4464.

XX  
XX

SQ Sequence 202 AA;

Query Match 53.6%; Score 52; DB 22; Length 202;  
Best Local Similarity 60.0%; Pred. No. 0.64;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KPLDKFGNIYDYHYE 15  
DB 41 qpdkfgelydl npe 55  
:  
:|:||||||| : |

RESULT 5  
AAG82008  
ID AAG82008 standard; Protein; 312 AA.  
XX AC AAG82008;  
XX AC  
DT 03-SEP-2001 (first entry)  
XX DE S. epidermidis open reading frame protein sequence SEQ ID NO:1110.  
XX KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;  
XX KW vaccination; endocarditis.  
XX OS Staphylococcus epidermidis.  
XX PN WO200134809-A2.  
XX PD 17-MAY-2001.  
XX PF 09-NOV-2000; 2000WO-US30782.  
XX PR 09-NOV-1999; 99US-0164258.  
XX PA (GLAX ) GLAXO GROUP LTD.  
PI Kimmerly WJ;  
XX WPI; 2001-316495/33.  
XX N-PSDB; AAH52858.  
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
XX useful for vaccinating against infections, e.g. endocarditis -  
XX Claim 18; Page 322; 2188pp; English.  
XX  
XX  
XX  
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.  
CC (I) and (II) can have antibacterial activity and therefore can be used  
CC in vaccination. The nucleic acids (I) may be used to produce the  
CC S. epidermidis polypeptides (II) via the production of vectors  
CC containing them which are used to produce hosts cells which express the  
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
CC used to vaccinate subjects and to raise antibodies against the bacteria.  
CC The polypeptides may also be used to assay for other inhibitors of their  
CC activity and therefore identify compounds that may be used for the  
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
CC AAH55090 represent specifically claimed s. epidermidis genomic DNA  
CC polynucleotide sequences from the present invention. AAH55091 to  
CC AAH55098 represent oligonucleotide sequences and primers which are used  
CC in the exemplification of the present invention.  
CC N.B. The present invention specifically claims all the polynucleotide  
CC sequences given in the sequence listing of the present specification,

CC however the sequence listing only goes up to SEQ ID NO:4454 so even  
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,  
CC no sequences are present for SEQ ID NO:4455 to 4464.

XX SQ Sequence 312 AA;

Query Match 53.6%; Score 52; DB 22; Length 312;  
Best Local Similarity 60.0%; Pred. No. 1;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDHYE 15  
:|:||||| ||| :|  
Db 41 qpdkdfgeiydlope 55

RESULT 6

AAAY35228  
ID AAY35228 standard; Protein; 262 AA.

XX AC AAY35228;

XX DT 13-SEP-1999 (first entry)

XX DE Amino acid sequence of a Chlamydia pneumoniae protein.

XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
KW vaccine; neutralising epitope.

XX OS Chlamydia pneumoniae.

XX PN WO9927105-A2.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-IB01890.

XX PR 04-NOV-1998; 98US-0107078.

XX PR 21-NOV-1997; 97FR-0014673.

XX PA (GEST ) GENSET.

XX PI Griffais R;

XX DR WPI; 1999-357842/30.

XX PT Genome sequence of Chlamydia pneumoniae

XX PS Page 1074; Disclosure; 1912pp; English.

XX CC AAY34584-Y35879 represent the proteins encoded by all the open reading  
frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.  
XX CC C. pneumoniae causes respiratory disease such as pneumonia and  
XX CC bronchitis and is thought to be a contributing factor in heart  
XX CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
XX CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
XX CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in  
XX CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
XX CC nucleotide sequences can also be used as immunogenic compositions,  
XX CC especially where the vector directs the expression of a neutralising  
XX CC epitope of C. pneumoniae.

XX SQ Sequence 262 AA;

Query Match 49.5%; Score 48; DB 20; Length 262;  
Best Local Similarity 57.1%; Pred. No. 3.9;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDHY 14  
:|:||||| ||  
Db 233 rpldkgfngirghy 246

RESULT 7

AAAY71539  
ID AAY71539 standard; Protein; 89 AA.

XX AC AAY71539;

XX DT 12-OCT-2000 (first entry)

XX DE Soybean POP2/CAFL1 transcription factor encoded by expressed sequence tag.

XX KW Soybean; POP2 transcription factor; CAFL1 transcription factor; ADH2 gene;  
KW glucose-repressible alcohol dehydrogenase gene; cell metabolism;  
KW transgenic plant; herbicide; expressed sequence tag; EST.

XX OS Glycine max.

XX PN WO200032782-A2.

XX PD 08-JUN-2000.

XX PF 01-DEC-1999; 99WO-US28322.

XX PR 02-DEC-1998; 98US-0110585.

XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Allen SM, Weng Z, Helentjaris TG;

XX DR WPI; 2000-412328/35.

XX DR N-PSDB; AAD01331.

XX PT New nucleic acids encoding transcription factor in plants and seeds,  
XX PT useful for producing transgenic plants, antibodies and selecting a  
XX PT polynucleotide that affects transcription factor polypeptide expression

XX PS Disclosure; Page 45; 46pp; English.

XX CC The present sequence is a CAFL1 (also known as POP2) transcription factor  
XX CC encoded by an expressed sequence tag from srl.pk0126.b4 clone isolated  
XX CC from soybean root cDNA library srl. CAFL1 protein has been shown to bind  
XX CC to another transcription factor CCR4 and regulate the expression of  
XX CC glucose-repressible alcohol dehydrogenase (ADH2) gene. The present  
XX CC sequence is useful for producing transgenic plants with altered  
XX CC levels of POP2/CAFL1 transcription factor to control expression of  
XX CC various genes or transgenes. The CAFL1 protein may also be used  
XX CC to design or identify inhibitors of cell metabolism that may be  
XX CC useful as herbicides.

XX SQ Sequence 89 AA;

Query Match 45.4%; Score 44; DB 21; Length 89;  
Best Local Similarity 46.7%; Pred. No. 5.4;  
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDHYE 15  
:|:||||| ||| :|  
Db 46 rpvgnfknindynyq 60

RESULT 8

AAAY71534  
ID AAY71534 standard; Protein; 309 AA.

XX AC AAY71534;

XX DT 12-OCT-2000 (first entry)

XX DE Soybean POP2/CAFL1 transcription factor.

XX Soybean; POP2 transcription factor; CAF1 transcription factor; ADH2 gene;  
KW glucose-repressible alcohol dehydrogenase gene; cell metabolism;  
KW transgenic plant; herbicide.  
XX  
OS Glycine max.  
XX  
XX WO200032782-A2.  
PN  
XX  
XX 08-JUN-2000.  
PD  
XX  
XX 01-DEC-1999; 99WO-US28322.  
PF  
XX  
XX 02-DEC-1998; 98US-0110585.  
PR  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
XX Allen SM, Weng Z, Helentjaris TG;  
PI  
XX WPI; 2000-412328/35.  
DR N-PSDB; AAD01326.  
DR  
XX  
XX New nucleic acids encoding transcription factor in plants and seeds,  
PT useful for producing transgenic plants, antibodies and selecting a  
PT polynucleotide that affects transcription factor polypeptide expression  
PT  
XX  
XX Claim 10; Page 38-39; 46pp; English.  
PS  
XX The present sequence is a CAF1 (also known as POP2)  
CC transcription factor from srl.pk0126.b4 clone isolated from soybean  
CC root cDNA library srl. CAF1 protein has been shown to bind  
CC to another transcription factor CCR4 and regulate the expression of  
CC glucose-repressible alcohol dehydrogenase (ADH2) gene. The present  
CC sequence is useful for producing transgenic plants with altered  
CC levels of POP2/CAF1 transcription factor to control expression of  
CC various genes or transgenes. The CAF1 protein may also be used  
CC to design or identify inhibitors of cell metabolism that may be  
CC useful as herbicides.  
XX  
XX Sequence 309 AA;  
SQ  
Query Match 45.4%; Score 44; DB 21; Length 309;  
Best Local Similarity 46.7%; Pred. No. 21;  
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
QY 1 KPLDRFGNIYDYHYE 15  
DB : : : : :  
85 rpvgnfknldynyq 99  
RESULT 9  
AAP91951  
ID AAP91951 standard; protein; 319 AA.  
XX  
AC AAP91951;  
XX  
DT 12-FEB-1990 (first entry)  
XX  
DE Polypeptide with lipase activity.  
XX  
KW Lipase.  
XX  
OS Chromobacterium viscosum var paralipoliticum.  
XX  
XX DE3908131-A.  
PN  
XX  
PD 05-OCT-1989.  
XX  
XX 13-MAR-1989; 89DE-3908131.  
PF  
XX

PR 14-MAR-1988; 88JP-0058376.  
XX  
XX (TOXN ) TOYO JOZO CO. LTD.  
XX  
XX Sagai H, Ohta H, Imamura S;  
XX  
XX WPI; 1989-293865/41.  
DR N-PSDB; AAN91331.  
DR  
XX  
XX New polypeptide with lipase activity - used for prodn. of fatty acids  
PT from triglyceride(s), etc.  
XX  
XX Claim 1; page 12; 16pp; German.  
PS  
XX The polypeptide has good stability and high lipase activity. It is used  
CC for prodn. of fatty acids from triglycerides, as a reagent for  
CC quantitative analysis of troglycerides, and as a catalyst for  
CC transesterifying fats and oils.  
XX  
XX Sequence 319 AA;  
SQ  
Query Match 44.3%; Score 43; DB 10; Length 319;  
Best Local Similarity 63.6%; Pred. No. 32;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 4 DKFGNIYDYHY 14  
DB : : : : :  
21 dkfanvvdwy 31  
RESULT 10  
AAW09624  
ID AAW09624 standard; Protein; 319 AA.  
XX  
AC AAW09624;  
XX  
DT 02-MAY-1997 (first entry)  
XX  
XX Pseudomonas glumae lipase modified to have reverse stereoselectivity.  
DE  
XX Pseudomonas; Chromobacterium; modification; variant; optical isomer;  
KW stereoselectivity; drug preparation; 1,4-dihydropyridine;  
KW substitution.  
XX  
OS Pseudomonas glumae.  
XX  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 221  
FT /label= substitution  
FT /note= "wild-type Phe replaced with Leu,  
FT important substitution for producing  
FT a lipase of reverse stereoselectivity"  
XX  
XX JP08256767-A.  
PN  
XX  
XX 08-OCT-1996.  
PD  
XX  
XX 22-JAN-1996; 96JP-0028640.  
PF  
XX  
XX 25-JAN-1995; 95JP-0030093.  
PR  
XX  
XX (AMAN ) AMANO PHARM KK.  
PA  
XX  
XX WPI; 1996-500357/50.  
DR  
XX  
XX Lipase of modified stereo:selectivity - useful for prepn. of  
PT optically active isomer 1,4-di:hydro:pyridine cpd., as intermediate  
PT for prepn. of drugs  
XX  
XX Disclosure; Page 9; 10pp; Japanese.  
PS  
XX  
XX AAW09620-W09625 are modified lipase sequences derived from Pseudomonas





CC sequences are based upon the wild-type P. glumae sequence disclosed  
CC in EP407225-A (UNILEVER PLC). The lipase variants may contain one  
CC or more of the amino acid substitutions indicated in the features  
CC table. Using these variants it was shown that lipases can be  
CC modified in a such a way that interaction with the substrate can be  
CC improved without forming large hydrophobic areas on the modified  
CC lipase surface which allow aggregation of lipase molecules. the  
CC lipase variants show improved in-the-wash lipolytic activity and may  
CC be used in enzymatic detergent compsns.

SQ Sequence 319 AA;

Query Match 44.3%; Score 43; DB 17; Length 319;

Best Local Similarity 63.6%; Pred. No. 32;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DKFGNIYDHY 14

||| | : |||

DB 21 dkfanvvdwy 31

RESULT 13

AAR88010

ID AAR88010 standard; protein: 319 AA.

XX

AC AAR88010;

XX

DT 02-AUG-1996 (first entry)

XX

DE Mature Pseudomonas glumae lipase T129Y variant.

XX

KW Mutant; lipase; enzymatic detergent; substitution; variant; improved;  
KW hydrophobicity.

XX

OS Pseudomonas glumae.

XX

FH Key Location/Qualifiers

FT Misc-difference 129

FT /label= substitution

FT /note= "Thr to Tyr, in wild-type sequence a Thr residue  
FT is present at position 129 of mature P. glumae  
FT lipase"

FT Misc-difference 23

FT /note= "possible site for Phe to Arg substitution"

FT Misc-difference 134

FT /note= "possible site for Leu to Arg substitution"

FT Misc-difference 148

FT /note= "possible site for Thr to Val substitution"

FT Misc-difference 233

FT /note= "possible site for Thr to Arg substitution"

FT Misc-difference 234

FT /note= "possible site for Leu to Arg substitution"

FT Misc-difference 239

FT /note= "possible site for Val to Phe substitution"

FT Misc-difference 240

FT /note= "possible site for Thr to Leu substitution"

FT Misc-difference 282

FT /note= "possible site for His to Arg substitution"

FT Misc-difference 292

FT /note= "possible site for Leu to Phe substitution"

XX W0535381-A1.

XX

PD 28-DEC-1995.

XX

PF 15-JUN-1995; 95WO-EP02349.

XX

PR 20-JUN-1994; 94EP-0201761.

XX

PA (UNIL ) UNILEVER NV.

PA (UNIL ) UNILEVER PLC.

XX

PI De Vlieg J, Frenken LGJ, Peters H, Suerbaum HMU;  
PI Verrips CT;

XX

DR WPI; 1996-058418/06.

XX

XX Pseudomonas lipase variant with increased surface hydrophobicity -  
XX PT has improved in-the-wash performance, useful in enzymatic detergent  
XX Claim 12; Page -; 33pp; English.

XX

CC AAR88009-R88018 are protein variants of Pseudomonas glumae lipase. The  
CC sequences are based upon the wild-type P. glumae sequence disclosed  
CC in EP407225-A (UNILEVER PLC). The lipase variants may contain one  
CC or more of the amino acid substitutions indicated in the features  
CC table. Using these variants it was shown that lipases can be  
CC modified in a such a way that interaction with the substrate can be  
CC improved without forming large hydrophobic areas on the modified  
CC lipase surface which allow aggregation of lipase molecules. the  
CC lipase variants show improved in-the-wash lipolytic activity and may  
CC be used in enzymatic detergent compsns.

SQ Sequence 319 AA;

Query Match 44.3%; Score 43; DB 17; Length 319;

Best Local Similarity 63.6%; Pred. No. 32;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DKFGNIYDHY 14

||| | : |||

DB 21 dkfanvvdwy 31

RESULT 14

AAR88011

ID AAR88011 standard; protein: 319 AA.

XX

AC AAR88011;

XX

DT 02-AUG-1996 (first entry)

XX

DE Mature Pseudomonas glumae lipase L134R variant.

XX

KW Mutant; lipase; enzymatic detergent; substitution; variant; improved;  
KW hydrophobicity.

XX

OS Pseudomonas glumae.

XX

FH Key Location/Qualifiers

FT Misc-difference 134

FT /label= substitution

FT /note= "Leu to Arg, in wild-type sequence a Leu residue  
FT is present at position 134 of mature P. glumae  
FT lipase"

FT Misc-difference 23

FT /note= "possible site for Phe to Arg substitution"

FT Misc-difference 129

FT /note= "possible site for Thr to Tyr substitution"

FT Misc-difference 148

FT /note= "possible site for Thr to Val substitution"

FT Misc-difference 233

FT /note= "possible site for Thr to Arg substitution"

FT Misc-difference 234

FT /note= "possible site for Leu to Arg substitution"

FT Misc-difference 239

FT /note= "possible site for Val to Phe substitution"

FT Misc-difference 240

FT /note= "possible site for Thr to Leu substitution"

FT Misc-difference 282

FT /note= "possible site for His to Arg substitution"

FT Misc-difference 292

FT /note= "possible site for Leu to Phe substitution"

XX

```
PN W09535381-AL.
XX
PD 28-DEC-1995.
XX
XX 15-JUN-1995; 95WO-EP02349.
XX
XX 20-JUN-1994; 94EP-0201761.
XX
XX (UNIL ) UNILEVER NV.
XX (UNIL ) UNILEVER PLC.
XX
XX De Vlieg J, Frenken LGJ, Peters H, Suerbaum HMU;
XX Verrips CT;
XX
XX WPI; 1996-058418/06.
XX
XX Pseudomonas lipase variant with increased surface hydrophobicity -
XX has improved in-the-wash performance, useful in enzymatic detergent
XX
XX Claim 12; Page -: 33pp; English.
XX
XX AAR88009-R88018 are protein variants of Pseudomonas glumae lipase. The
XX sequences are based upon the wild-type P. glumae sequence disclosed
XX in EP407225-A (UNILEVER PLC). The lipase variants may contain one
XX or more of the amino acid substitutions indicated in the features
XX table. Using these variants it was shown that lipases can be
XX modified in a such a way that interaction with the substrate can be
XX improved without forming large hydrophobic areas on the modified
XX lipase surface which allow aggregation of lipase molecules. the
XX lipase variants show improved in-the-wash lipolytic activity and may
XX be used in enzymatic detergent compans.
XX
XX Sequence 319 AA;

Query Match 44.3%; Score 43; DB 17; Length 319;
Best Local Similarity 63.6%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DKFGNIYDHY 14
   ||| | : || |
Db 21 dkfanvvdwy 31

RESULT 15
AAR88012
ID AAR88012 standard; protein; 319 AA.
XX
XX AAR88012;
XX
XX 02-AUG-1996 (first entry)
XX
XX Mature Pseudomonas glumae lipase T148V variant.
XX
XX Mutant; lipase; enzymatic detergent; substitution; variant; improved;
XX hydrophobicity.
XX
XX Pseudomonas glumae.
XX
XX Key Location/Qualifiers
XX
XX Misc-difference 148
XX /label- substitution
XX /note= "Thr to Val, in wild-type sequence a Thr residue
XX is present at position 148 of mature P. glumae
XX lipase"
XX
XX Misc-difference 23
XX /note= "possible site for Phe to Arg substitution"
XX
XX Misc-difference 129
XX /note= "possible site for Thr to Tyr substitution"
XX
XX Misc-difference 134
XX /note= "possible site for Leu to Arg substitution"
XX
XX Misc-difference 233
XX /note= "possible site for Thr to Arg substitution"
```

```
FT Misc-difference 234
FT /note= "possible site for Leu to Arg substitution"
FT
FT Misc-difference 239
FT /note= "possible site for Val to Phe substitution"
FT
FT Misc-difference 240
FT /note= "possible site for Thr to Leu substitution"
FT
FT Misc-difference 282
FT /note= "possible site for His to Arg substitution"
FT
FT Misc-difference 292
FT /note= "possible site for Leu to Phe substitution"
XX
XX W09535381-AL.
PN
XX
XX 28-DEC-1995.
XX
XX 15-JUN-1995; 95WO-EP02349.
XX
XX 20-JUN-1994; 94EP-0201761.
XX
XX (UNIL ) UNILEVER NV.
XX (UNIL ) UNILEVER PLC.
XX
XX De Vlieg J, Frenken LGJ, Peters H, Suerbaum HMU;
XX Verrips CT;
XX
XX WPI; 1996-058418/06.
XX
XX Pseudomonas lipase variant with increased surface hydrophobicity -
XX has improved in-the-wash performance, useful in enzymatic detergent
XX
XX Claim 12; Page -: 33pp; English.
XX
XX AAR88009-R88018 are protein variants of Pseudomonas glumae lipase. The
XX sequences are based upon the wild-type P. glumae sequence disclosed
XX in EP407225-A (UNILEVER PLC). The lipase variants may contain one
XX or more of the amino acid substitutions indicated in the features
XX table. Using these variants it was shown that lipases can be
XX modified in a such a way that interaction with the substrate can be
XX improved without forming large hydrophobic areas on the modified
XX lipase surface which allow aggregation of lipase molecules. the
XX lipase variants show improved in-the-wash lipolytic activity and may
XX be used in enzymatic detergent compans.
XX
XX Sequence 319 AA;

Query Match 44.3%; Score 43; DB 17; Length 319;
Best Local Similarity 63.6%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DKFGNIYDHY 14
   ||| | : || |
Db 21 dkfanvvdwy 31
```

Search completed: January 29, 2002, 10:21:41  
Job time: 417 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:00 ; Search time 133.18 Seconds  
(without alignments)  
2.704 Million cell updates/sec

Title: US-09-763-397A-3

Perfect score: 97

Sequence: 1 KPLDKFGNIYDYHYEH 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	63.9	20	1 US-08-120-225-1	Sequence 1, Appli
2	43	44.3	358	1 US-08-034-650-10	Sequence 10, Appl
3	43	44.3	358	1 US-08-449-015-10	Sequence 10, Appl
4	39.5	40.7	1294	2 US-08-819-288-3	Sequence 3, Appli
5	39.5	40.7	1321	1 US-08-261-822A-3	Sequence 3, Appli
6	39.5	40.7	1321	5 PCT-US95-07744A-3	Sequence 3, Appli
7	39	40.2	242	1 US-08-596-864-3	Sequence 3, Appli
8	39	40.2	242	2 US-08-989-667-3	Sequence 3, Appli
9	39	40.2	374	1 US-08-625-876-2	Sequence 2, Appli
10	39	40.2	374	1 US-08-875-876-4	Sequence 4, Appli
11	39	40.2	374	1 US-08-625-876-6	Sequence 6, Appli
12	39	40.2	379	1 US-07-863-169A-5	Sequence 5, Appli
13	39	40.2	379	2 US-08-424-268-8	Sequence 8, Appli
14	39	40.2	379	2 US-08-429-964-5	Sequence 5, Appli
15	39	40.2	379	3 US-07-935-087-5	Sequence 5, Appli
16	39	40.2	379	5 PCT-US93-08062-5	Sequence 5, Appli
17	39	40.2	379	5 PCT-US93-10442-8	Sequence 8, Appli
18	39	40.2	819	2 US-08-424-268-20	Sequence 20, Appl
19	39	40.2	819	5 PCT-US93-10442-20	Sequence 20, Appl
20	38	39.2	350	2 US-08-495-695B-33	Sequence 33, Appl
21	38	39.2	375	1 US-08-176-412-2	Sequence 2, Appli
22	38	39.2	375	2 US-08-555-268A-2	Sequence 2, Appli
23	38	39.2	375	2 US-08-495-695B-2	Sequence 2, Appli
24	38	39.2	375	5 PCT-US94-14436-2	Sequence 2, Appli
25	38	39.2	432	2 US-08-472-172-6	Sequence 6, Appli
26	38	39.2	455	2 US-08-472-172-4	Sequence 4, Appli
27	38	39.2	460	1 US-08-689-974-5	Sequence 5, Appli

28	38	39.2	460	3	US-09-058-376-5	Sequence 5, Appli
29	38	39.2	485	1	US-07-991-867B-42	Sequence 42, Appl
30	38	39.2	485	2	US-08-544-332-42	Sequence 42, Appl
31	38	39.2	618	4	US-08-961-083-72	Sequence 72, Appl
32	38	39.2	1456	1	US-08-026-138E-8	Sequence 8, Appli
33	38	39.2	1482	1	US-08-026-138E-2	Sequence 2, Appli
34	38	39.2	1484	2	US-08-231-193A-56	Sequence 56, Appl
35	38	39.2	1484	2	US-08-486-273A-56	Sequence 56, Appl
36	38	39.2	1484	3	US-08-940-086A-56	Sequence 56, Appl
37	38	39.2	3135	1	US-08-323-170B-2	Sequence 2, Appli
38	37	38.1	21	1	US-08-244-701B-42	Sequence 42, Appl
39	37	38.1	25	1	US-08-244-701B-11	Sequence 11, Appl
40	37	38.1	31	1	US-08-244-701B-8	Sequence 8, Appli
41	37	38.1	255	1	US-08-242-188-4	Sequence 4, Appli
42	37	38.1	255	1	US-08-509-261A-4	Sequence 4, Appli
43	37	38.1	255	1	US-08-660-626-10	Sequence 10, Appl
44	37	38.1	255	1	US-08-692-892-4	Sequence 4, Appli
45	37	38.1	255	2	US-08-713-939A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-08-120-225-1  
; Sequence 1, Application US/08120225  
; Patent No. 5502168  
; GENERAL INFORMATION:  
; APPLICANT: Kumar, Nirbhay  
; TITLE OF INVENTION: NOVEL CONTINUOUS AND CROSS-REACTING  
; TITLE OF INVENTION: EPITOPE FOR HUMAN MALARIA VACCINE  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L. Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/120,225  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/779,494  
; FILING DATE: 24-OCT-1991  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-120-225-1

Query Match 63.9%; Score 62; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.00048;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KPLDKFGNIYD 11

Db 10 KPLDKFGNIYD 20

RESULT 2

US-08-034-650-10

; Sequence 10, Application US/08034650  
; Patent No. 5641671  
; GENERAL INFORMATION:  
; APPLICANT: BOS, Jannetje W.  
; APPLICANT: FRENKEN, Leon G.  
; APPLICANT: VERRIPS, Cornelis T.  
; APPLICANT: VISSER, Christiaan  
; TITLE OF INVENTION: PRODUCTION OF ACTIVE PSEUDOMONAS GLUMAE  
; TITLE OF INVENTION: LIPASE IN HOMOLOGOUS OR HETEROLOGOUS HOSTS  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L. Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036-5601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/034,650  
; FILING DATE: 03-JUL-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: US 07/727,235  
; APPLICATION DATA:  
; NAME: Kokulis, Paul N.  
; REGISTRATION NUMBER: 16,773  
; REFERENCE/DOCKET NUMBER: PNK/5970/91731  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 358 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-034-650-10

Query Match 44.3%; Score 43; DB 1; Length 358;  
Best Local Similarity 63.6%; Pred. No. 15;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DKFGNIYDHY 14  
||| | | |  
Db 60 DKFANVDYVY 70

RESULT 3  
US-08-449-015-10  
; Sequence 10, Application US/08449015  
; Patent No. 580409  
; GENERAL INFORMATION:  
; APPLICANT: BOS, Jannetje W.  
; APPLICANT: FRENKEN, Leon G.  
; APPLICANT: VERRIPS, Cornelis T.  
; APPLICANT: VISSER, Christiaan  
; TITLE OF INVENTION: PRODUCTION OF ACTIVE PSEUDOMONAS GLUMAE  
; TITLE OF INVENTION: LIPASE IN HOMOLOGOUS OR HETEROLOGOUS HOSTS  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L. Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA

; ZIP: 20036-5601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/449,015  
; FILING DATE: 24-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: US 07/727,235  
; APPLICATION DATA:  
; NAME: Kokulis, Paul N.  
; REGISTRATION NUMBER: 16,773  
; REFERENCE/DOCKET NUMBER: PNK/5970/91731  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 358 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-449-015-10

Query Match 44.3%; Score 43; DB 1; Length 358;  
Best Local Similarity 63.6%; Pred. No. 15;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DKFGNIYDHY 14  
||| | | |  
Db 60 DKFANVDYVY 70

RESULT 4  
US-08-819-288-3  
; Sequence 3, Application US/08819288  
; Patent No. 5955652  
; GENERAL INFORMATION:  
; APPLICANT: Ecker, Joseph  
; APPLICANT: Alonso, Jose  
; TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE  
; TITLE OF INVENTION: AND PATHOGENS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5955652ris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/819,288  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Beardell, Lori Y.  
; REGISTRATION NUMBER: 34,293  
; REFERENCE/DOCKET NUMBER: UPN-2949  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1294 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-819-288-3

Query Match 40.7%; Score 39.5; DB 2; Length 1294;  
Best Local Similarity 58.3%; Pred. No. 2.3e+02;  
Matches 7; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 3 LDKF-GNIYDYH 13  
||:| |:|:|:|  
Db 678 LDEFWGHLYDFH 689

RESULT 5  
US-08-261-822A-3  
; Sequence 3, Application US/08261822A  
; Patent No. 5650553  
; GENERAL INFORMATION:  
; APPLICANT: Ecker, Joseph R. et al.  
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene  
; NUMBER OF SEQUENCES: 82  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553rls  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/261,822A  
; FILING DATE: 17-JUN-1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Beardell, Lori Y.  
; REGISTRATION NUMBER: 34,293  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1321 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-261-822A-3

Query Match 40.7%; Score 39.5; DB 1; Length 1321;  
Best Local Similarity 58.3%; Pred. No. 2.4e+02;  
Matches 7; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 3 LDKF-GNIYDYH 13  
||:| |:|:|:|  
Db 678 LDEFWGHLYDFH 689

RESULT 6  
PCT-US95-07744A-3  
; Sequence 3, Application PC/TUS9507744A  
; GENERAL INFORMATION:  
; APPLICANT: Trustees of The University of Pennsylvania  
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene  
; NUMBER OF SEQUENCES: 82  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/07744A  
; FILING DATE: 15-JUNE-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER: 08/261,822  
; FILING DATE: June 17, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Beardell, Lori Y.  
; REGISTRATION NUMBER: 34,293  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1321 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
PCT-US95-07744A-3

Query Match 40.7%; Score 39.5; DB 5; Length 1321;  
Best Local Similarity 58.3%; Pred. No. 2.4e+02;  
Matches 7; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 3 LDKF-GNIYDYH 13  
||:| |:|:|:|  
Db 678 LDEFWGHLYDFH 689

RESULT 7  
US-08-596-864-3  
; Sequence 3, Application US/08596864  
; Patent No. 5731183  
; GENERAL INFORMATION:  
; APPLICANT: KOBAYASHI, KATSUNORI  
; APPLICANT: YAMANAKA, SHIGERU  
; APPLICANT: MIWA, KIYOSHI  
; APPLICANT: SUZUKI, SHUNICHI  
; APPLICANT: ETO, YUZURU  
; APPLICANT: TANITA, YUKO  
; APPLICANT: YOKOZEKI, KENZO  
; APPLICANT: HASHIGUCHI, KENICHI  
; TITLE OF INVENTION: BACILLUS-DERIVED TRANSGLUTAMINASE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/596,864
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: JP 021963/1995
; FILING DATE: 09-FEB-1995
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 226316/1995
; FILING DATE: 04-SEP-1995
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 013072/1996
; FILING DATE: 29-JAN-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-786-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 242 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-596-864-3
```

```
Query Match 40.2%; Score 39; DB 1; Length 242;
Best Local Similarity 44.4%; Pred. No. 44;
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;
```

```
QY 4 DKFGN-----IYDVHYE 15
||| :||:|
Db 135 DKFNASFDRIILYDWHYE 152
```

```
RESULT 8
; Sequence 3, Application US/089898667
; Patent No. 5948662
; GENERAL INFORMATION:
; APPLICANT: KOBAYASHI, KATSUNORI
; APPLICANT: YAMANAKA, SHIGERU
; APPLICANT: MIWA, KIYOSHI
; APPLICANT: SUZUKI, SHUNICHI
; APPLICANT: ETO, YUZURO
; APPLICANT: TANITA, YUKO
; APPLICANT: YOKOZEKI, KENICHI
; TITLE OF INVENTION: BACILLUS-DERIVED TRANSLUTAMINASE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLEAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,667
```

```
;
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/596,864
; FILING DATE: 09-FEB-1996
; APPLICATION NUMBER: JP 021963/1995
; FILING DATE: 09-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 226316/1995
; FILING DATE: 04-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 013072/1996
; FILING DATE: 29-JAN-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-786-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 242 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-989-667-3
```

```
Query Match 40.2%; Score 39; DB 2; Length 242;
Best Local Similarity 44.4%; Pred. No. 44;
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;
```

```
QY 4 DKFGN-----IYDVHYE 15
||| :||:|
Db 135 DKFNASFDRIILYDWHYE 152
```

```
RESULT 9
; Sequence 2, Application US/08625876
; Patent No. 5656471
; GENERAL INFORMATION:
; APPLICANT: MINAGAWA, HIROTAKE
; APPLICANT: NAKAYAMA, NO. 5656471Iiyuki
; APPLICANT: NAKAMOTO, SHINYA
; TITLE OF INVENTION: LACTATE OXIDASE WITH AN IMPROVED THERMAL
; STABILITY AND GENE OF THE SAME
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,876
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-95947
; FILING DATE: 30-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-146186
; FILING DATE: 13-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J
```



REGISTRATION NUMBER: 32,925  
REFERENCE/DOCKET NUMBER: PF-1612  
TELEPHONE: (703) 521-2297  
TELEFAX: (703) 685-0573  
TELEX: 248425 EMBON  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 374 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-625-876-2

Query Match 40.2%; Score 39; DB 1; Length 374;  
Best Local Similarity 57.1%; Pred. No. 71;  
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDHY 14  
| | | | | | | |  
Db 361 KGLDLFDNPGYGY 374

RESULT 10  
US-08-625-876-4  
; Sequence 4, Application US/08625876  
; Patent No. 5656471  
; GENERAL INFORMATION:  
; APPLICANT: MINAGAWA, Hirotaka  
; APPLICANT: NAKAYAMA, No. 5656471Iiyuki  
; APPLICANT: NAKAMOTO, Shinya  
; TITLE OF INVENTION: LACTATE OXIDASE WITH AN IMPROVED THERMAL  
; TITLE OF INVENTION: STABILITY AND GENE OF THE SAME  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: YOUNG & THOMPSON  
; STREET: 745 South 23rd Street  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/625,876  
FILING DATE: 01-APR-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-95947  
FILING DATE: 30-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-146186  
FILING DATE: 13-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: PATCH, Andrew J  
REGISTRATION NUMBER: 32,925  
REFERENCE/DOCKET NUMBER: PF-1612  
TELEPHONE: (703) 521-2297  
TELEFAX: (703) 685-0573  
TELEX: 248425 EMBON  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 374 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-625-876-4

Query Match 40.2%; Score 39; DB 1; Length 374;  
Best Local Similarity 57.1%; Pred. No. 71;  
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDHY 14  
| | | | | | | |  
Db 361 KGLDLFDNPGYGY 374

RESULT 11  
US-08-625-876-6  
; Sequence 6, Application US/08625876  
; Patent No. 5656471  
; GENERAL INFORMATION:  
; APPLICANT: MINAGAWA, Hirotaka  
; APPLICANT: NAKAYAMA, No. 5656471Iiyuki  
; APPLICANT: NAKAMOTO, Shinya  
; TITLE OF INVENTION: LACTATE OXIDASE WITH AN IMPROVED THERMAL  
; TITLE OF INVENTION: STABILITY AND GENE OF THE SAME  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: YOUNG & THOMPSON  
; STREET: 745 South 23rd Street  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/625,876  
FILING DATE: 01-APR-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-95947  
FILING DATE: 30-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-146186  
FILING DATE: 13-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: PATCH, Andrew J  
REGISTRATION NUMBER: 32,925  
REFERENCE/DOCKET NUMBER: PF-1612  
TELEPHONE: (703) 521-2297  
TELEFAX: (703) 685-0573  
TELEX: 248425 EMBON  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 374 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-625-876-6

Query Match 40.2%; Score 39; DB 1; Length 374;  
Best Local Similarity 57.1%; Pred. No. 71;  
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KPLDKFGNIYDHY 14  
| | | | | | | |  
Db 361 KGLDLFDNPGYGY 374

RESULT 12  
US-07-863-169A-5  
; Sequence 5, Application US/07863169A  
; Patent No. 5420245

```

; GENERAL INFORMATION:
; APPLICANT: Brown, Michael S.
; APPLICANT: Goldstein, Joseph L.
; APPLICANT: Reiss, Yuval
; TITLE OF INVENTION: Tetrapeptide-Based Inhibitors of Farnesyl
; TITLE OF INVENTION: Transferase
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/863,169A
; FILING DATE: 03-APR-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/822,011
; FILING DATE: 19-JAN-1992
; CLASSIFICATION: 530
; APPLICATION NUMBER: US 07/937,893
; FILING DATE: 18-APR-1991
; CLASSIFICATION: 530
; APPLICATION NUMBER: US 615,715
; FILING DATE: 20-NOV-1990
; CLASSIFICATION: 530
; APPLICATION NUMBER: US 510,706
; FILING DATE: 18-APR-1990
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:297/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-863-169A-5

```

Query Match 40.28; Score 39; DB 1; Length 379;  
 Best Local Similarity 66.7%; Pred. No. 72;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 DKFGNIYD 12  
 III :|||  
 DB 94 DKFRVDYD 102

```

; RESULT 13
; US-08-424-268-8
; Sequence 8, Application US/08424268
; Patent No. 5821118
; GENERAL INFORMATION:
; APPLICANT: Omer, Charles A
; APPLICANT: Diehl, Ronald E
; APPLICANT: Gibbs, Jackson B
; APPLICANT: Kohl, Nancy E
; TITLE OF INVENTION: Assay for Inhibitors of Farnesyl-Protein
; TITLE OF INVENTION: Transferase

```

```

; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O.Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Mac
; OPERATING SYSTEM: System 7.5.3
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,268
; FILING DATE: 4/24/95
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Muthard, David A
; REGISTRATION NUMBER: 35,297
; REFERENCE/DOCKET NUMBER: 18858PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3903
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; US-08-424-268-8

```

Query Match 40.2%; Score 39; DB 2; Length 379;  
 Best Local Similarity 66.7%; Pred No. 72;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 DKFGNIYD 12  
 III :|||  
 DB 94 DKFRVDYD 102

```

; RESULT 14
; US-08-429-964-5
; Sequence 5, Application US/08429964
; Patent No. 5962243
; GENERAL INFORMATION:
; APPLICANT: BROWN, MICHAEL S.
; APPLICANT: GOLDSTEIN, JOSEPH L.
; APPLICANT: REISS, YUVAL
; APPLICANT: JAMES, GUY L.
; TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
; TITLE OF INVENTION: TRANSFERASE INHIBITORS
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/429,964
; FILING DATE: 27-APR-1995
; CLASSIFICATION: 435

```

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,625  
FILING DATE: 16-FEB-1993  
CLASSIFICATION: 435  
APPLICATION NUMBER: US 07/822,011  
FILING DATE: ABANDONED  
CLASSIFICATION: 435  
APPLICATION NUMBER: PCT/US/91/02650  
FILING DATE: 18-APR-1991  
CLASSIFICATION: 435  
APPLICATION NUMBER: US 07/615,715  
FILING DATE: 20-NOV-1990  
CLASSIFICATION: 435  
APPLICATION NUMBER: US 07/510,706  
FILING DATE: 18-APR-1990 (ABANDONED)  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: PARKER, DAVID L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UTSD:432/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 379 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-429-964-5

Query Match 40.2%; Score 39; DB 2; Length 379;  
Best Local Similarity 66.7%; Pred. No. 72;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Oy 4 DKFGNIYDY 12  
||| :|||  
Db 94 DKFRDVIDY 102

RESULT 15  
US-07-935-087-5  
Sequence 5, Application US/07935087  
Patent No. 6083917  
GENERAL INFORMATION:  
APPLICANT: BROWN, MICHAEL S.  
APPLICANT: GOLDSTEIN, JOSEPH L.  
APPLICANT: REISS, YUVAL  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR  
TITLE OF INVENTION: THE IDENTIFICATION,  
TITLE OF INVENTION: CHARACTERIZATION,  
TITLE OF INVENTION: AND INHIBITION OF FARNESYL  
TITLE OF INVENTION: PROTEIN TRANSFERASE  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE & DURKEE  
STREET: P.O. BOX 4433  
CITY: HOUSTON  
STATE: TEXAS  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1 (converted to ASCII-DOS)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,087  
FILING DATE: 19920824  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/822,011  
FILING DATE: 01/16/92  
ATTORNEY/AGENT INFORMATION:  
NAME: PARKER, DAVID L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UTSD:269/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512-320-7200  
TELEFAX: 512-474-7577  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 379 amino acid residues  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-935-087-5

Query Match 40.2%; Score 39; DB 3; Length 379;  
Best Local Similarity 66.7%; Pred. No. 72;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Oy 4 DKFGNIYDY 12  
||| :|||  
Db 94 DKFRDVIDY 102

Search completed: January 29, 2002, 10:24:01  
Job time: 507 sec

us-09-763-397a-3.rai

Mon Feb 4 15:23:47 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:36 : Search time 80.65 seconds  
(without alignments)  
7.274 Million cell updates/sec

Title: US-09-763-397A-3

Perfect score: 97

Sequence: 1 KPLDKFGNIYDYHYEH 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	46.4	392	1 PORA_THEME	O05651 thermotoga
2	44.5	45.9	316	1 TALIA_ECOLI	P78258 escherichia
3	44.5	45.9	507	1 CATYA_PICAN	P30263 pichia angu
4	44	45.4	591	1 CYB2_YEAST	P00175 saccharomyc
5	44	45.4	867	1 VL96_IRV1	P22856 tipula irid
6	44	45.4	1072	1 ITA6_CHICK	P26007 gallus gall
7	44	45.4	1534	1 MTDM_ARATH	P34881 arabidopsis
8	43	44.3	358	1 LIP_PSEGL	Q05489 pseudomonas
9	43	44.3	540	1 TOPI_AQUAE	O64148 mesocricetu
10	42	43.3	395	1 NH10_CAEEL	O66893 aquifex aeo
11	42	43.3	1683	1 YJ09_YEAST	P41999 caenorhabdi
12	41.5	42.8	759	1 ARV2_CALVI	P47054 saccharomyc
13	41	42.3	273	1 RFA2_YEAST	P28514 calliphora
14	41	42.3	503	1 CP3A_MESAU	P26754 saccharomyc
15	40	41.2	241	1 KDNM_BPT4	O64148 mesocricetu
16	40	41.2	251	1 YC52_METJA	Q58648 methanococc
17	40	41.2	358	1 MTH2_HAEP4	P15446 haemophilus
18	40	41.2	364	1 LEU3_AQUAE	O66607 aquifex aeo
19	40	41.2	429	1 GUNA_BUTFI	P22541 butyrivibri
20	40	41.2	433	1 POP2_YEAST	P39008 saccharomyc
21	40	41.2	491	1 YC75_YEAST	P25381 saccharomyc
22	40	41.2	503	1 CP3E_CAVPO	Q64417 cavia porce
23	40	41.2	503	1 CP3F_CAVPO	Q64406 cavia porce
24	40	41.2	503	1 CP3H_CAVPO	Q64409 cavia porce
25	40	41.2	678	1 NCPR_SCHPO	P36587 schizosacch
26	40	41.2	787	1 YD5A_SCHPO	Q10313 schizosacch
27	40	41.2	876	1 RPOB_NPVAC	P41452 autographa
28	40	41.2	1115	1 DP3A_BACSU	Q34523 bacillus su
29	39	40.2	245	1 TGL_BACSU	P40746 bacillus su
30	39	40.2	286	1 LPVC_CHLTR	O84538 chlamydia t
31	39	40.2	379	1 PFTA_HUMAN	P49354 homo sapien
32	39	40.2	398	1 YWJE_BACSU	P45865 bacillus su
33	39	40.2	544	1 MATK_MAIZE	P48190 zea mays (m

RESULT 1	PORA_THEME	STANDARD;	PRT;	392 AA.
ID	AC	O05651;		
DT	15-DEC-1998	(Rel. 37, Created)		
DT	15-DEC-1998	(Rel. 37, Last sequence update)		
DT	20-AUG-2001	(Rel. 40, Last annotation update)		
DE	DE	PYRUVATE SYNTHASE SUBUNIT PORA (EC 1.2.7.1) (PYRUVATE OXIDOREDUCTASE ALPHA CHAIN) (POR) (PYRUVIC-FERREDOXIN OXIDOREDUCTASE ALPHA SUBUNIT).		
GN	PORA OR TM0017.			
OS	Thermotoga maritima.			
OC	Bacteria; Thermotogales; Thermotoga.			
OX	NCBI_TaxID=2336;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-43.			
RC	STRAIN=MSB8 / DSM 3109;			
RX	MEDLINE=96125254; PubMed=8550425;			
RA	Kletzin A., Adams M.W.W.A.;			
RT	"Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from Pyrococcus furiosus and pyruvate ferredoxin oxidoreductase from Thermotoga maritima.";			
RL	J. Bacteriol. 178:248-257(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MSB8 / DSM 3109;			
RX	MEDLINE=99287316; PubMed=10360571;			
RA	Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M., Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D., Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.C., Fraser C.M.;			
RT	"Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima.";			
RL	Nature 399:323-329(1999).			
RN	[3]			
RP	SEQUENCE OF 1-43, AND CHARACTERIZATION.			
RC	STRAIN=MSB8 / DSM 3109;			
RX	MEDLINE=94137707; PubMed=8305426;			
RA	Blamey J.M., Adams M.W.W.A.;			
RT	"Characterization of an ancestral type of pyruvate ferredoxin oxidoreductase from the hyperthermophilic bacterium, Thermotoga maritima.";			
RL	Biochemistry 33:1000-1007(1994).			
CC	-1- CATALYTIC ACTIVITY: PYRUVATE + COA + OXIDIZED FERREDOXIN = ACETYL-COA + CO(2) + REDUCED FERREDOXIN.			
CC	-1- SUBUNIT: HETEROTETRAMER OF ONE ALPHA, ONE BETA, ONE DELTA AND ONE GAMMA CHAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			

Q47861 erwinia her  
P42846 saccharomyc  
P43573 saccharomyc  
O51124 borrelia bu  
P25178 helicobacte  
O92m3 helicobacte  
O51044 borrelia bu  
P36185 entamoeba h  
O01958 entamoeba h  
P79217 oryctolagus  
P50391 homo sapien  
O62762 rattus norv

#### ALIGNMENTS

```
-----
CC EMBL; X85171; CAA59457.1; ALT_INIT.
DR EMBL; AE001690; AAD35111.1; ALT_INIT.
DR TIGR; TW0017; -.
DR InterPro; IPR002880; POR_N.
DR Pfam; PF01855; POR_N; 1.
KW Oxidoreductase; Complete proteome.
FT DOMAIN 72 79 POLY-ALA.
SQ SEQUENCE 392 AA; 44111 MW; ACEF0E34C9EC1818 CRC64;

Query Match 46.4%; Score 45; DB 1; Length 392;
Best Local Similarity 53.3%; Pred. No. 8.2; Indels 4; Gaps 1;
Matches 8; Conservative 3; Mismatches 0;

QY 2 PLDRFGNIYDYHYEH 16
Db 214 PLD-----LYDYFHEH 224

RESULT 2
TALA_ECOLI STANDARD; PRT; 316 AA.
ID TALA_ECOLI
AC P78258; PR0218;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TRANSALDOLASE A (EC 2.2.1.2).
GN TALA OR B2464 OR Z3720 OR ECS3326.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Iida A., Teshiba S., Mizobuchi K.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
```

```
-----
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509552;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobé T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [6]
RP PRESENCE OF TWO TRANSALDOLASES IN E. COLI.
RA Sprenger G.A.;
RL Unpublished observations (JUN-1993).
CC -!- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF
CC METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.
CC -!- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
CC 3-PHOSPHATE = D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.
CC -!- PATHWAY: NONOXIDATIVE PART OF THE PENTOSE-PHOSPHATE PATHWAY.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC
CC EMBL; D13159; BAA21821.1; -.
CC EMBL; AE000333; AAC75517.1; -.
DR EMBL; D90875; BAA16339.1; -.
DR EMBL; AE005475; AAG57573.1; -.
DR EMBL; AP002561; BAB36749.1; -.
DR HSP; P30148; IUCW.
DR EcoGene; EG11797; talA.
DR InterPro; IPR001585; Transaldolase.
DR Pfam; PF00923; Transaldolase; 1.
DR PROSITE; PS00958; TRANSALDOLASE_2; 1.
DR PROSITE; PS01054; TRANSALDOLASE_1; 1.
KW Transferase; Pentose shunt; Multigene family; Complete proteome.
FT ACT_SITE 131 131 BY SIMILARITY
FT SEQUENCE 316 AA; 35658 MW; 990B00ED7937CF19 CRC64;
SQ
Query Match 45.9%; Score 44.5; DB 1; Length 316;
Best Local Similarity 33.3%; Pred. No. 7.8;
Matches 9; Conservative 4; Mismatches 3; Indels 11; Gaps 1;

QY 1 KPLDKF-----GNIYDYHYEH 16
Db 189 KPMDPYVVEEDPGVKVRNIYDYKQH 215

RESULT 3
CATA_PICAN STANDARD; PRT; 507 AA.
ID CATA_PICAN
AC P30263;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN PXP9 OR PXP-9.
OS Pichia angusta (Yeast) (Hansenula polymorpha).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4905;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN-ATCC 34438;  
RX MEDLINE=92299073; PubMed=1607006;  
RA Didion T., Roggenkamp R.O.;  
RT "Targeting signal of the peroxisomal catalase in the methylotrophic  
RL yeast *Hansenula polymorpha*.";  
FEBS Lett. 303:113-116(1992).  
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
CC PEROXIDE.  
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
CC -!- COFACTOR: HEME GROUP.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- SUBCELLULAR LOCATION: PEROXISOMAL.  
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X56501; CAA39856.1; -;  
DR HSSP: P00432; 7CAT.  
DR InterPro: IPR002226; Catalase.  
DR Pfam: PF00199; catalase; 1.  
DR PRINTS: PR00067; CATALASE.  
DR ProDom: PD000510; Catalase; 1.  
DR PROSITE: PS00342; MICROBODIES\_CTER; 1.  
DR PROSITE: PS00437; CATALASE\_1; 1.  
DR PROSITE: PS00438; CATALASE\_2; 1.  
DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
KW Peroxisome. 65 65  
FT ACT\_SITE 65 138 BY SIMILARITY.  
FT ACT\_SITE 138 138 BY SIMILARITY.  
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).  
FT SITE 505 507 MICROBODY TARGETING SIGNAL (POTENTIAL).  
SQ SEQUENCE 507 AA; 57849 MW; 3536ED0A49539CC3 CRC64;  
  
Query Match 45.9%; Score 44.5; DB 1; Length 507;  
Best Local Similarity 47.1%; Pred. NO. 13;  
Matches 8; Conservative 4; Mismatches 4; Indels 1; Gaps 1;  
  
QY 1 KPLDKF-GNIYDYHYEH 16  
DB 415 KPDEYTGVPYHYEH 431  
  
RESULT 4  
CYB2\_YEAST  
ID CYB2\_YEAST STANDARD; PRT; 591 AA.  
AC P00175;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE CYTOCHROME B2 PRECURSOR (EC 1.1.1.23) (L-LACTATE DEHYDROGENASE  
DE [CYTOCHROME]) (L-LACTATE FERRICYTOCHROME C OXIDOREDUCTASE) (L-LCR).  
GN CYB2 OR YML054C OR YN9558.08C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RX MEDLINE=86135959; PubMed=3004948;  
RA Guillard B.;  
RT "Structure, expression and regulation of a nuclear gene encoding a  
RT mitochondrial protein: the yeast L(+)-lactate cytochrome c  
RT oxidoreductase (cytochrome b2).";  
RL EMBO J. 4:3265-3272(1985).

RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Devlin K., Churcher C., Barrell B.G., Rajandream M.A.;  
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 81-394.  
RX MEDLINE=84132029; PubMed=6365548;  
RA Ghir R., Becam A.-M., Lederer F.;  
RT "Primary structure of flavocytochrome b2 from baker's yeast.  
RT Purification by reverse-phase high-pressure liquid chromatography and  
RT sequencing of fragment alpha cyanogen bromide peptides.";  
RL Eur. J. Biochem. 139:59-74(1984).  
RN [4]  
RP SEQUENCE OF 395-591.  
RX MEDLINE=86030284; PubMed=3902473;  
RA Lederer F., Cortial S., Becam A.M., Haumont P.-Y., Perez L.;  
RT "Complete amino acid sequence of flavocytochrome b2 from baker's  
RT yeast.";  
RL Eur. J. Biochem. 152:419-428(1985).  
RN [5]  
RP SEQUENCE OF 81-94.  
RX MEDLINE=75156546; PubMed=165435;  
RA Guillard B., Lederer F., Jacq C.;  
RT "More similarity between bakers' yeast L-(+)-lactate dehydrogenase and  
RT liver microsomal cytochrome B5.";  
RL Nature 255:422-423(1975).  
RN [6]  
RP SEQUENCE OF 88-183.  
RX MEDLINE=76206228; PubMed=776230;  
RA Guillard B., Lederer F.;  
RT "Complete amino acid sequence of the heme-binding core in bakers'  
RT yeast cytochrome b2 (L-(+)-lactate dehydrogenase).";  
RL Biochimie 58:305-316(1976).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
RX MEDLINE=90230315; PubMed=2329585;  
RA Xia Z.-X., Mathews F.S.;  
RT "Molecular structure of flavocytochrome b2 at 2.4-A resolution.";  
RL J. Mol. Biol. 212:837-863(1990).  
CC -!- CATALYTIC ACTIVITY: L-LACTATE + 2 FERRICYTOCHROME C = PYRUVATE +  
CC 2 FERROCITYCHROME C.  
CC -!- COFACTOR: THIS ENZYME BINDS FMN AND PROTOHEME IX PROSTHETIC  
CC GROUPS.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INTERMEMBRANE SPACE.  
CC -!- INDUCTION: BY L-LACTATE. INDUCED DURING RESPIRATORY ADAPTATION.  
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTOCHROME  
CC B5 FAMILY.  
CC -!- SIMILARITY: TO SPINACH GLYCOLATE OXIDASE (33% IDENTITY), TO  
CC P. PUTIDA S-MANDELATE DEHYDROGENASE, AND TO M. SMEGMATIS LACTATE  
CC 2-MONOOXYGENASE.  
CC -!- DATABASE: NAME=worthington enzyme manual;  
CC WWW="http://www.worthington-biochem.com/manual/L/YLDH.html".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X03215; CAA26959.1; -;  
DR EMBL: Z46729; CAA86721.1; -;  
DR PIR: A24583; CBBY2.  
DR PIR: A23095; A23095.  
DR PDB: 1FCB; 15-JUL-92.  
DR PDB: 1LTD; 31-AUG-94.  
DR PDB: 1LCO; 15-SEP-95.  
DR PDB: 1LDC; 10-JUL-95.  
DR SGD: S0004518; CYB2.  
DR InterPro: IPR001199; Cyt\_B5.

DR InterPro: IPR003009; FMN\_enzyme.  
DR InterPro: IPR000262; FMN\_hydroxy\_acid\_dh.  
DR Pfam: PF01070; FMN\_dh; 1.  
DR Pfam: PF00173; heme\_1; 1.  
DR PROSITE: PS00191; CYTOCHROME\_B5\_1; 1.  
DR PROSITE: PS0255; CYTOCHROME\_B5\_2; 1.  
DR PROSITE: PS00557; FMN\_HYDROXY\_ACID\_DH; 1.  
KW Electron transport; respiratory chain; Oxidoreductase; Flavoprotein;  
FMN; Heme; Transit peptide; Mitochondrion; 3D-structure.  
TRANSIT 1 80 MITOCHONDRION.  
CHAIN 81 591 CYTOCHROME B2.  
DOMAIN 81 176 CYTOCHROME DOMAIN (CORE).  
DOMAIN 177 190 HINGE.  
DOMAIN 191 566 FMN-BINDING DOMAIN.  
DOMAIN 567 591 TAIL (WRAPS AROUND THE MOLECULAR 4-FOLD  
AXIS).  
FT 123 123 HEME LIGAND (BY SIMILARITY).  
FT 146 146 HEME LIGAND (BY SIMILARITY).  
FT 453 453 REMOVES THE SUBSTRATE ALPHA-PROTON AS THE  
ACT\_SITE 453 453 FIRST STEP IN CATALYSIS.  
FT 456 456 SUBSTRATE BINDING.  
FT 165 165 Q -> E (IN REF. 3 AND 6).  
FT 466 466 E -> Q (IN REF. 4).  
FT 513 513 R -> E (IN REF. 4).  
FT 91 91  
FT 91 91  
FT 93 98  
FT 102 104  
FT 106 109  
FT 110 111  
FT 111 111  
FT 112 115  
FT 117 119  
FT 120 122  
FT 124 125  
FT 128 131  
FT 132 133  
FT 137 137  
FT 139 145  
FT 148 149  
FT 150 154  
FT 157 159  
FT 160 164  
FT 173 173  
FT 179 180  
FT 184 184  
FT 185 191  
FT 199 201  
FT 206 215  
FT 218 225  
FT 229 230  
FT 232 243  
FT 245 246  
FT 258 258  
FT 261 263  
FT 264 265  
FT 266 268  
FT 272 274  
FT 280 283  
FT 285 288  
FT 289 297  
FT 305 308  
FT 309 310  
FT 315 320  
FT 321 321  
FT 321 321  
FT 329 333  
FT 339 351  
FT 352 353  
FT 357 360  
FT 376 376  
FT 399 401  
FT 408 409  
FT 412 420  
FT 421 421  
FT 421 431  
FT 426 431

FT TURN 435 436  
FT HELIX 437 442  
FT TURN 443 444  
FT STRAND 447 450  
FT TURN 453 456  
FT TURN 459 460  
FT STRAND 461 461  
FT HELIX 464 476  
FT TURN 477 479  
FT TURN 481 483  
FT STRAND 485 489  
FT HELIX 495 503  
FT TURN 504 505  
FT STRAND 508 511  
FT HELIX 513 545  
FT TURN 546 546  
FT HELIX 550 552  
FT HELIX 555 557  
FT STRAND 558 559  
FT TURN 561 564  
FT STRAND 568 569  
FT HELIX 574 579  
SQ SEQUENCE 591 AA; 65539 MW; DBADA0751B3C5B83 CRC64;

Query Match 45.4%; Score 44; DB 1; Length 591;  
Best Local Similarity 53.8%; Pred. NO. 18;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 PLDKFGNIYDHYH 14  
Db 198 PLDNIINLYDFEY 210

RESULT 5  
VL96\_IRV1 STANDARD; PRT; 867 AA.  
ID VL96\_IRV1  
AC P22856;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE L96 PROTEIN.  
GN L96.  
OS Tipula iridescent virus (TIV) (Insect iridescent virus type 1).  
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.  
OX NCBI\_TaxID=10490;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91078646; PubMed=1701750;  
RA Home W.A., Tajbakhsh S., Seligy V.L.;  
RT "Molecular cloning and characterization of a late Tipula iridescent  
virus gene."  
RL Gene 94:243-248(1990).  
CC -!- FUNCTION: MAY BE INVOLVED IN TIV GENOMIC DNA PACKAGING IN A  
MANNER RELATED TO THE GAG POLYPROTEINS OF THE MAMMALIAN VIRUSES.  
CC -!- SIMILARITY: CONTAINS 1 OTU DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M62953; AAA47919.1;  
CC PIR; JH0225; JH0225.  
DR InterPro: IPR003323; OTU.  
DR Pfam; PF02338; OTU; 1.  
KW Repeat; DNA packaging; DNA-binding.  
FT DOMAIN 606 745 OTU.  
SQ SEQUENCE 867 AA; 96011 MW; F19DBDE8FE5CA103 CRC64;



```
Query Match 45.4%; Score 44; DB 1; Length 867;
Best Local Similarity 64.3%; Pred. No. 27;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KPLDKFGNIYDYH 14
    ||| | | | | |
DB 726 KPLPLGFIIDYH 739

RESULT 6
ITAG6_CHICK STANDARD; PRT; 1072 AA.
AC P26007;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE INTEGRIN ALPHA-6 PRECURSOR (VLA-6).
GN ITGA6.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91185416; PubMed=1826298;
RA "Laminin receptors in the retina: sequence analysis of the chick
RT integrin alpha 6 subunit. Evidence for transcriptional and
RT posttranslational regulation.";
RL J. Cell Biol. 113:405-416(1991).
CC -!- FUNCTION: INTEGRIN ALPHA-6/BETA-1 IS A RECEPTOR FOR LAMININ ON
CC PLATELETS. INTEGRIN ALPHA-6/BETA-4 IS A RECEPTOR FOR LAMININ IN
CC EPITHELIAL CELLS. IT PLAYS A CRITICAL STRUCTURAL ROLE IN THE
CC HEMIDESMOSOME.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
CC DISULFIDE BOND. ALPHA-6 ASSOCIATES WITH EITHER BETA-1 OR BETA-4.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- DEVELOPMENTAL STAGE: ALPHA-6 LEVELS DECREASE WITH AGE.
CC -!- PTM: PHOSPHORYLATED IN VIVO.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56559; CAA39909.1; .
DR PIR; A38457; A38457.
DR HSSP; P11215; 1A8X.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 5.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Extracellular matrix; Cytoskeleton; Phosphorylation; Repeat;
KW Calcium.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 1072 INTEGRIN ALPHA-6.
FT CHAIN 19 898 INTEGRIN ALPHA-6 HEAVY CHAIN (POTENTIAL).
FT CHAIN 902 1072 INTEGRIN ALPHA-6 LIGHT CHAIN (POTENTIAL).
FT DOMAIN 19 1010 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1011 1036 POTENTIAL.
FT DOMAIN 1037 1072 CYTOPLASMIC (POTENTIAL).
```

```
FT REPEAT 34 72 FG-GAP 1.
FT REPEAT 107 139 FG-GAP 2.
FT REPEAT 179 211 FG-GAP 3.
FT REPEAT 250 286 FG-GAP 4.
FT REPEAT 308 346 FG-GAP 5.
FT REPEAT 369 405 FG-GAP 6.
FT REPEAT 427 467 FG-GAP 7.
FT CA_BIND 318 326 POTENTIAL.
FT CA_BIND 380 388 POTENTIAL.
FT CA_BIND 438 446 POTENTIAL.
FT SITE 1039 1043 GFPR MOTIF.
FT DISULFID 79 88 BY SIMILARITY.
FT DISULFID 125 148 BY SIMILARITY.
FT DISULFID 169 182 BY SIMILARITY.
FT DISULFID 498 557 BY SIMILARITY.
FT DISULFID 625 631 BY SIMILARITY.
FT DISULFID 725 736 BY SIMILARITY.
FT DISULFID 880 927 INTERCHAIN (BY SIMILARITY).
FT DISULFID 933 938 BY SIMILARITY.
FT CARBOHYD 71 71 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 730 730 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 780 780 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 957 957 N-LINKED (GLCNAC. .) (POTENTIAL).
FT MOD_RES 1070 1070 PHOSPHORYLATION (BY CAM-KINASE II)
                                (POTENTIAL).
SO SEQUENCE 1072 AA; 119199 MW; 6DE2B99ECD8B3E44 CRC64;

Query Match 45.4%; Score 44; DB 1; Length 1072;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 PLDKFGNIYDYH 13
    | | | | | |
DB 394 PYDGFGRVYIH 405

RESULT 7
ITDM_ARATH STANDARD; PRT; 1534 AA.
AC P34881;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DNA (CYTOSINE-5)-METHYLTRANSFERASE ATHI (EC 2.1.1.37) (DNA
DE METHYLTRANSFERASE ATHI) (DNA METASE ATHI) (M.ATHI).
GN ATHI.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=93281384; PubMed=8389441;
RA Finnegan E.J., Dennis E.S.;
RT "Isolation and identification by sequence homology of a putative
RT cytosine methyltransferase from Arabidopsis thaliana.";
RL Nucleic Acids Res. 21:2383-2388(1993).
CC -!- FUNCTION: METHYLATES CG RESIDUES.
CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA -> S-ADENOSYL-L-
CC HOMOCYSTEINE + DNA CONTAINING 5-METHYLCYTOSINE.
CC -!- SIMILARITY: LOW, TO OTHER EUKARYOTIC DNA METASE.
CC -!- SIMILARITY: SOME TO BACTERIAL RESTRICTION SYSTEMS
CC METHYLTRANSFERASES.
CC -!- SIMILARITY: CONTAINS 2 BAH DOMAINS.
```

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
CC EMBL: L10692; AAA32829.1; --  
DR REBASE: 2839; M.Athi.  
DR InterPro: IPR001025; BAH.  
DR InterPro: IPR001525; C5\_DNA\_meth.  
DR Pfam: PF01426; BAH; 2.  
DR Pfam: PF00145; DNA\_methylase; 3.  
DR PRINTS: PR00105; C5METHYLASE.  
DR SMART: SM00439; BAH; 2.  
DR PROSITE: PS00094; C5\_MTASE\_1; 1.  
DR PROSITE: PS00095; C5\_MTASE\_2; 1.  
DR Transferase: Methyltransferase; DNA-binding.  
KW ACT\_SITE 1198 1198 BY SIMILARITY.  
FT ACT\_SITE 1534 AA; 172430 MW; 23FC944AA7074C5A CRC64;  
SQ SEQUENCE 1534 AA; 172430 MW; 23FC944AA7074C5A CRC64;  
  
Query Match 45.4%; Score 44; DB 1; Length 1534;  
Best Local Similarity 46.7%; Pred. No. 50;  
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
  
Qy 1 KPLDKFGNIYDHYE 15  
|| : |||||  
Db 166 KPAASRYKVDYFYE 180  
  
RESULT 8  
LIP\_PSEGL LIP\_PSEGL STANDARD; PRT; 358 AA.  
AC Q05489;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE LIPASE PRECURSOR (EC 3.1.1.3) (TRIACYLGLYCEROL LIPASE).  
GN LIPA.  
OS Pseudomonas glumae, and Chromobacterium viscosum.  
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;  
OC Burkholderia.  
OX NCBI\_TaxID=337, 42739;  
RN [1]  
RP SEQUENCE FROM N.A., SEQUENCE OF 40-61, AND MUTAGENESIS.  
RC SPECIES=P.glumae; STRAIN=PGI / CBS 322.89;  
RA MEDLINE=93119130; PubMed=1476423;  
RX Frenken L.G.J., Egmond M.R., Batenburg A.M., Bos J.W., Visser C.,  
RA Verrips C.T.;  
RT "Cloning of the Pseudomonas glumae lipase gene and determination of  
RT the active site residues."  
RL Appl. Environ. Microbiol. 58:3787-3791(1992).  
RN [2]  
RP SEQUENCE OF 40-54, AND CHARACTERIZATION.  
RC SPECIES=C.viscosum;  
RX MEDLINE=95506500; PubMed=7786905;  
RA Taipa M.A., Liebeton K., Costa J.V., Cabral J.M.S., Jaeger K.-E.;  
RT "Lipase from Chromobacterium viscosum: biochemical characterization  
RT indicating homology to the lipase from Pseudomonas glumae."  
RL Biochim. Biophys. Acta 1256:396-402(1995).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).  
RC SPECIES=P.glumae;  
RX MEDLINE=94009622; PubMed=8405390;  
RA Noble M.E.M., Cleasby A., Johnson L.N., Egmond M.R., Frenken L.G.J.;  
RT "The crystal structure of triacylglycerol lipase from Pseudomonas  
RT glumae reveals a partially redundant catalytic aspartate."  
RL FEBS Lett. 331:123-128(1993).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).

SPECIES=C.viscosum; STRAIN=ATCC 6918;  
MEDLINE=96275656; PubMed=8683577;  
RA Lang D., Hofmann B., Haalck L., Hecht H.-J., Spener F., Schmid R.D.,  
RA Schomburg D.;  
RT "Crystal structure of a bacterial lipase from Chromobacterium  
RT viscosum ATCC 6918 refined at 1.6-A resolution.";  
RL J. Mol. Biol. 259:704-717(1996).  
CC -!- FUNCTION: HYDROLYSIS OF TRIGLYCERIDES.  
CC -!- CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL +  
CC A FATTY ACID ANION.  
CC -!- COFACTOR: REQUIRES CALCIUM.  
CC -!- SUBUNIT: MONOMER.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- SIMILARITY: STRONG TO OTHER PSEUDOMONAS LIPASES.  
CC -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,  
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
CC EMBL: X70354; CAA49812.1; --  
DR EMBL: A16323; CAA01279.1; --  
DR EMBL: A32021; CAA02073.1; --  
DR PIR: A48952; A48952.  
DR PIR: S37291; S37291.  
DR PDB: 1TAH; 31-MAY-94.  
DR InterPro: IPR000073; Abhydrolase.  
DR InterPro: IPR000379; Est\_lip\_thioest\_actsite.  
DR InterPro: IPR000734; Lipase.  
DR Pfam: PF00561; abhydrolase; 1.  
DR PROSITE: PS00120; LIPASE\_SER; 1.  
KW Hydrolase; Lipid degradation; Signal; Calcium; 3D-structure.  
FT SIGNAL 1 39  
FT CHAIN 40 358 LIPASE.  
FT ACT\_SITE 126 126 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 302 302 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 324 324 CHARGE RELAY SYSTEM.  
FT DISULFID 229 308  
FT MUTAGEN 54 54 H->A: NO LOSS OF ACTIVITY.  
FT MUTAGEN 126 126 S->A: COMPLETE LOSS OF ACTIVITY.  
FT MUTAGEN 160 160 D->E: NO LOSS OF ACTIVITY.  
FT MUTAGEN 160 160 D->A: NO LOSS OF ACTIVITY.  
FT MUTAGEN 280 280 D->E: NO LOSS OF ACTIVITY.  
FT MUTAGEN 280 280 D->A: COMPLETE LOSS OF ACTIVITY.  
FT MUTAGEN 302 302 D->E: NO LOSS OF ACTIVITY.  
FT MUTAGEN 302 302 D->A: 75% LOSS OF ACTIVITY.  
FT MUTAGEN 324 324 H->A: COMPLETE LOSS OF ACTIVITY.  
FT CONFLICT 40 40 A -> W (IN REF. 2).  
SQ SEQUENCE 358 AA; 36928 MW; FE7B5D7A22EC6B4B CRC64;  
  
Query Match 44.3%; Score 43; DB 1; Length 358;  
Best Local Similarity 63.6%; Pred. No. 15;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 4 DKFGNIYDHY 14  
||| | : |||  
Db 60 DKFANVDYDY 70  
  
RESULT 9  
TOP1\_AQAE  
ID TOP1\_AQAE STANDARD; PRT; 540 AA.  
AC O66893;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DNA TOPOISOMERASE I (EC 5.99.1.2) (OMEGA-PROTEIN) (RELAXING ENZYME)



DR EMBL: Z49314; CAA89330.1; -  
DR SGD: S0003576; YJ039C.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 228 248 POTENTIAL.  
FT TRANSMEM 268 288 POTENTIAL.  
FT TRANSMEM 312 332 POTENTIAL.  
FT TRANSMEM 430 450 POTENTIAL.  
FT TRANSMEM 600 620 POTENTIAL.  
FT TRANSMEM 634 654 POTENTIAL.  
FT TRANSMEM 747 767 POTENTIAL.  
FT TRANSMEM 783 803 POTENTIAL.  
FT TRANSMEM 912 932 POTENTIAL.  
FT TRANSMEM 1022 1042 POTENTIAL.  
FT TRANSMEM 1295 1315 POTENTIAL.  
FT TRANSMEM 1343 1363 POTENTIAL.  
FT TRANSMEM 1506 1526 POTENTIAL.  
FT TRANSMEM 1595 1615 POTENTIAL.  
SQ SEQUENCE 1683 AA; 191534 MW; FD2D61FDEA545104 CRC64;

Query Match 43.38; Score 42; DB 1; Length 1683;  
Best Local Similarity 61.58; Pred. No. 1.1e+02;  
Matches 8; Conservative 3; Mismatches 0; Indels 2; Gaps 1;

QY 3 LDKF--GNIYDYH 13  
I::: I::: I:::  
DB 1228 LNKYCSGNIYDFH 1240

RESULT 12  
ARY2\_CALVI  
ID ARY2\_CALVI STANDARD; PRT; 759 AA.  
AC P28514;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DE 01-NOV-1997 (Rel. 35, Last annotation update)  
DE ARYLPHORIN SUBUNIT C223 PRECURSOR.  
OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
OX NCBI\_TaxID=7373;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fat body;  
RX MEDLINE=91282793; PubMed=1711849;  
RA Naumann U., Scheller K.;  
RT "Complete cDNA and gene sequence of the developmentally regulated arylphorin of Calliphora vicina and its homology to insect hemolymph proteins and arthropod hemocyanins.";  
RL Biochem. Biophys. Res. Commun. 177:963-972(1991).  
CC -!- FUNCTION: ARYLPHORIN IS A LARVAL STORAGE PROTEIN (LSP) WHICH MAY SERVE AS A STORAGE PROTEIN USED PRIMARILY AS A SOURCE OF AROMATIC AMINO ACIDS FOR PROTEIN SYNTHESIS DURING METAMORPHOSIS. IT IS A CONSTITUENT OF THE SCLEROTIZING SYSTEM OF THE CUTICLE, AND SERVES AS A CARRIER FOR ECDYSTEROID HORMONE.  
CC -!- SUBUNIT: HETEROHEXAMER.  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
CC -!- TISSUE SPECIFICITY: FAT BODY.  
CC -!- SIMILARITY: TO OTHER ARYLPHORINS AND TO ARTHROPOD HEMOCYANINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M76479; -; NOT ANNOTATED\_CDS.  
DR EMBL: X59390; CAA42033.1; -  
DR PIR: J01044; J01044.  
DR HSP: P04253; ILLI.

DR InterPro: IPR000896; Hemocyanin.  
DR Pfam: PF00372; hemocyanin; 1.  
DR PRINTS: PR00187; HAEMOCYANIN.  
DR PROSITE: PS00209; HEMOCYANIN\_1; FALSE\_NEG.  
DR PROSITE: PS00210; HEMOCYANIN\_2; 1.  
KW Signal; Storage protein; Glycoprotein; Multigene family.  
FT SIGNAL 1 15 POTENTIAL.  
FT CHAIN 16 759 ARYLPHORIN SUBUNIT C223.  
SQ SEQUENCE 759 AA; 92525 MW; 8B56CC7A2D71352D CRC64;

Query Match 42.88; Score 41.5; DB 1; Length 759;  
Best Local Similarity 53.88; Pred. No. 59;  
Matches 7; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 4 DKFGNIYDYHYEH 16  
I::: I::: I:::  
DB 748 EKFGH-FDNYHH 759

RESULT 13  
RFA2\_YEAST  
ID RFA2\_YEAST STANDARD; PRT; 273 AA.  
AC P26754; P38905;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DE 01-NOV-1997 (Rel. 35, Last annotation update)  
DE REPLICATION FACTOR-A PROTEIN 2 (RF-A) (DNA BINDING PROTEIN BUFL1)  
DE (REPLICATION PROTEIN A 36 KDA SUBUNIT).  
GN RFA2 OR BUFL1 OR YNL312W OR N0368.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 100-118; 150-172 AND 184-196.  
RC STRAIN=W303-1A;  
RX MEDLINE=91357474; PubMed=1885001;  
RA Brill S.J., Stillman R.;  
RT "Replication factor-A from Saccharomyces cerevisiae is encoded by three essential genes coordinately expressed at S phase.";  
RL Genes Dev. 5:1589-1600(1991).  
RN [2]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 5-16.  
RX MEDLINE=93361008; PubMed=8355713;  
RA Luche R.M., Smart W.C., Marion T., Tillman M., Sumrada R.A., Cooper T.G.;  
RT "Saccharomyces cerevisiae BUF protein binds to sequences participating in DNA replication in addition to those mediating transcriptional repression (URS1) and activation.";  
RL Mol. Cell. Biol. 13:5749-5761(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / FY1676;  
RX MEDLINE=96076632; PubMed=7502583;  
RA Maftah M., Nicaud J.-M., Levesque H., Gaillardin C.;  
RT "Sequencing analysis of a 24.7 Kb fragment of yeast chromosome XIV identifies six known genes, a new member of the hexose transporter family and ten new open reading frames.";  
RL Yeast 11:1077-1085(1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Maurer C.T.C., Urbanus J.H.M., Planta R.J.;  
RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases  
CC -!- FUNCTION: BINDS TO SINGLE-STRANDED SEQUENCES PARTICIPATING IN DNA REPLICATION IN ADDITION TO THOSE MEDIATING TRANSCRIPTIONAL REPRESSION (URS1) AND ACTIVATION (CAR1). STIMULATES THE ACTIVITY OF A COGNATE TOPISOMERASE I TO UNWIND TEMPLATE DNA CONTAINING THE T-AG AND DNA TOPOISOMERASE I TO UNWIND TEMPLATE DNA CONTAINING THE SIMIAN VIRUS 40 ORIGIN OF DNA REPLICATION.  
CC -!- SUBUNIT: HETEROPRIMER OF 69, 36, AND 13 KDA CHAINS. THE DNA-BINDING ACTIVITY MAY RESIDE EXCLUSIVELY ON THE 69 KDA SUBUNIT.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.

CC	TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,
CC	AND CARCINOGENS.
CC	-1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement/">http://www.isb-sib.ch/announcement/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL: M73992; ; NOT_ANNOTATED_CDS.
DR	EMBL: J79317; AAB35091.1; -
DR	InterPro: IPR001128; Cyt_P450.
DR	Fram: PF00067; P450; 1.
DR	PRINTS; PR00359; BP450.
DR	PRINTS; PR00385; P450.
DR	PRINTS; PR00463; EP450I.
DR	PRINTS; PR00464; EP450II.
DR	PRINTS; PR00465; EP450IV.
DR	PROSITE; PS00086; CYTOCHROME_P450; 1.
KW	Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW	Microsome; Endoplasmic reticulum.
FT	BINDING 442 442 HEME (BY SIMILARITY).
SO	SEQUENCE 503 AA; 57693 MW; D4D24FEE87FD7F51 CRC64;

Query Match	42.3%	Score 41:	DB 1:	Length 503:
Best Local Similarity	61.5%	Pred. No. 45:		
Matches 8:	Conservative	0:	Mismatches	5:
			Indels	0:
			Gaps	0:

```

Qy 1 KPLDKFGNIYDYH 13
    ||| || | ||
Db 42 KPLPFFGTILGYH 54

RESULT 15
KDNM_BPT4
ID KDNM_BPT4 STANDARD; PRT; 241 AA.
AC P04531;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DEOXYNUCLEOTIDE MONOPHOSPHATE KINASE (EC 2.7.4.13) (DNK).
GN 1.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=10665;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=86037230; PubMed=4057254;
RX Broida J., Abelson J.;
RA "Sequence organization and control of transcription in the
RT bacteriophage T4 trna region.";
RL J. Mol. Biol. 185:545-563(1985).
RN [2]
RN SEQUENCE OF 148-241 FROM N.A.
RP MEDLINE=82150240; PubMed=7063418;
RX Herrmann R.;
RA "Nucleotide sequence of the bacteriophage T4 gene 57 and a deduced
RT amino acid sequence.";
RL Nucleic Acids Res. 10:1105-1112(1982).
RN [3]
RN SEQUENCE OF 1-27 FROM N.A.
RP MEDLINE=89296504; PubMed=2740234;
RX Koch T., Lamm N., Rueger W.;
RA "Sequencing, cloning and overexpression of genes of bacteriophage T4
RT between map positions 74.325 and 77.184.";
RL Nucleic Acids Res. 17:4392-4392(1989).
RN [4]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RP

```

Mon Feb 4 15:23:48 2002

```

RX MEDLINE-96312945; PubMed-8670851;
RA Teplyakov A., Sebastiao P., Obmolova G., Perrakis A., Brush G.S.,
RA Bessman M.J., Wilson K.S.;
RT "Crystal structure of bacteriophage T4 deoxynucleotide kinase with
RL its substrates dGMP and ATP.";
RL ENBO J. 15:3487-3497(1996).
CC -1- FUNCTION: ACTS ON dGMP, dTMP AND 5-HYDROXYMETHYL-dCMP WHILE
CC EXCLUDING dCMP AND dAMP.
CC -1- CATALYTIC ACTIVITY: ATP + DEOXYNUCLEOSIDE PHOSPHATE = ADP +
CC DEOXYNUCLEOSIDE DIPHOSPHATE.
CC -1- SUBUNIT: HOMODIMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03016; CAA26800.1; -
DR EMBL; X14845; CAA32953.1; -
DR EMBL; J02516; AAA32497.1; -
DR PIR; A04309; KIBPD4.
DR PIR; S04612; S04612.
DR PDB; 1DEK; 11-JAN-97.
DR PDB; 1DEL; 11-JAN-97.
KW Transferase; Kinase; 3D-structure.
SQ SEQUENCE 241 AA; 27329 MW; 61AD2375CC94BDE8 CRC64;
-----
Query Match 41.2%; Score 40; DB 1; Length 241;
Best Local Similarity 63.6%; Pred. No. 30;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDKFGNIYDYH 13
   ||| : |||
Db 161 LDKFNSGYDYY 171

```

Search completed: January 29, 2002, 11:13:37  
Job time: 813 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:41 ; Search time 310.82 Seconds  
(without alignments)  
2.860 Million cell updates/sec

Title: US-09-763-397A-4

Perfect score: 69

Sequence: 1 NANPNANPNANP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_ll01:\*

1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT:\*

2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT:\*

3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT:\*

4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT:\*

5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT:\*

6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT:\*

7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT:\*

8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT:\*

9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT:\*

10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT:\*

11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT:\*

12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT:\*

13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT:\*

14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT:\*

15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT:\*

16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT:\*

17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT:\*

18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT:\*

19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT:\*

20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT:\*

21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT:\*

22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69	100.0	12	AA03092	Plasmodium falcipa
2	69	100.0	12	AA01783	Plasmodium falcipa
3	69	100.0	12	AA04981	Synthetic peptide
4	69	100.0	12	AA070038	Plasmodium falcipa
5	69	100.0	12	AA061556	Malaria circumspor
6	69	100.0	12	AA029731	Malaria B-cell ep
7	69	100.0	12	AA011790	Plasmodium falcipa
8	69	100.0	12	AA029734	Plasmodium falcipa
9	69	100.0	12	AA067599	Plasmodium falcipa
10	69	100.0	12	AA070280	Plasmodium falcipa
11	69	100.0	12	AA073753	Plasmodium falcipa

12	69	100.0	12	22	AA084362	Amino acid sequenc
13	69	100.0	12	22	AA049997	Tumour antigen MUC
14	69	100.0	13	7	AA060804	Sequence of an imm
15	69	100.0	13	15	AA049980	Synthetic peptide
16	69	100.0	13	16	AA090240	Malaria antigen fo
17	69	100.0	13	16	AA070040	fMAl protein C-te
18	69	100.0	14	10	AA090813	Non-immunogenic am
19	69	100.0	15	22	AA073754	Resin-bound Plasm
20	69	100.0	16	10	AA093496	3 full and 2 half
21	69	100.0	16	16	AA070039	fMAl protein N-te
22	69	100.0	16	20	AA036733	Amino acid sequenc
23	69	100.0	16	21	AA015523	P. falciparum circ
24	69	100.0	17	13	AA022907	Synthetic immunoge
25	69	100.0	19	14	AA038530	Plasmodium sporozo
26	69	100.0	20	19	AA069278	Malaria parasite p
27	69	100.0	24	22	AA031715	Synthetic NANP pep
28	69	100.0	27	21	AA080105	Amino acid sequenc
29	69	100.0	28	11	AA080362	Modified CDR3 of t
30	69	100.0	28	15	AA059808	Sequence of mutage
31	69	100.0	28	15	AA059873	Sequence of antigen
32	69	100.0	28	16	AA067668	CDR3 of antigenic
33	69	100.0	28	19	AA029733	Malaria TLB sequ
34	69	100.0	28	19	AA061558	Malaria circumspor
35	69	100.0	33	22	AA063516	A peptide which ma
36	69	100.0	36	21	AA091224	Modified MVF Th ep
37	69	100.0	36	21	AA091225	Modified MVF Th ep
38	69	100.0	38	9	AA082566	Anti-malarial pep
39	69	100.0	56	22	AA063514	A peptide which ma
40	69	100.0	59	22	AA063513	Vaccine against pl
41	69	100.0	61	9	AA082341	Immunogenic branch
42	69	100.0	63	12	AA014263	Immunogenic branch
43	69	100.0	64	12	AA014261	Immunogenic branch
44	69	100.0	65	12	AA014265	Immunogenic branch
45	69	100.0	65	12	AA014262	Immunogenic branch

## ALIGNMENTS

### RESULT 1

AA03092  
ID AA03092 standard; peptide; 12 AA.

AC AA03092;

DT 19-JUL-1990 (first entry)

DE Plasmodium falciparum immunodominant epitope.

KW Circumsporozoitic protein; immunodominant epitope; antimalaria vaccine.

OS Plasmodium falciparum.

PN EP353823-A.

PD 07-FEB-1990.

PF 28-JUL-1989; 89EP-0201985.

PR 05-AUG-1988; 88IT-0216688.

PA (ENTE ) ENRITERCHE SPA.

PI Verdini AS, Pessi A, DelGiulice G, Lambert PH;

DR WPI; 1990-038546/06.

PT Immunologically active peptide conjugate anti-malaria vaccine -  
comprising synthetic peptide corresponding to Plasmodium protein  
linked to tuberculin protein derivative.

PS Claim 3; page 5; 11pp; English.

Mon Feb 4 15:23:48 2002

CC The sequence NANP can be repeated 3-40 times. When  
 CC conjugated to a tuberculin protein derivative, the product  
 CC is useful as an anti-malaria vaccine and provides a high  
 CC antibody response.

XX Sequence 12 AA;  
 SQ

Query Match 100.0%; Score 69; DB 11; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
 | | | | | | | | | | | |  
 Db 1 nanpnanpnanp 12

RESULT 2

AAR41783  
 ID AAR41783 standard; protein; 12 AA.

XX AC AAR41783;

XX 25-MAR-1994 (first entry)

XX Plasmodium falciparum cs protein antigen.

XX Group A Streptococci; surface antigen; anchor region;  
 KW gram positive bacterium; M protein; fusion protein; immunogen;  
 KW malaria; circumsporozoite protein; vaccine.

XX Plasmodium falciparum.

XX WO9318163-A.

XX 16-SEP-1993.

XX 12-MAR-1993; 93WO-US02355.

XX 13-MAR-1992; 92US-0851082.

XX (UYRQ ) UNIV ROCKEFELLER.

XX Fischetti VA, Pozzi G, Schneewind O;

XX WPI; 1993-303477/38.

XX Gene encoding hybrid surface protein of gram positive bacteria -  
 PT useful for preparing vaccine compsn. for protecting animals from  
 PT bacterial infection

XX Disclosure; Page 53; 85pp; English.

XX Hybrid surface proteins are claimed in which an active  
 CC polypeptide (e.g. a surface antigen from a mammalian tumour cell,  
 CC sperm or an allergen, bacterium, virus, parasite or fungus) is  
 CC fused to an anchor region from a surface antigen normally expressed  
 CC on the cell surface of gram positive bacteria. The anchor segment  
 CC is pref. derived from a streptococcal M protein (see AAR41780).  
 CC Sequence AAR41783 is a peptide sequence from the cs protein of  
 CC P.falciparum (Nature 314:111-114) suitable for development of a  
 CC vaccine against malaria.

XX Sequence 12 AA;

Query Match 100.0%; Score 69; DB 14; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
 | | | | | | | | | | | |  
 Db 1 nanpnanpnanp 12

RESULT 3

AAR49981  
 ID AAR49981 standard; peptide; 12 AA.

XX AC AAR49981;

XX 10-OCT-1994 (first entry)

XX Synthetic peptide analogue of P. falciparum used as immunogen.

XX Peptide; retro; Inverso; retro-inverso; therapy; disease; vaccine;  
 KW immunogen; antibody; antibodies; assay; treatment; prophylaxis;  
 KW Plasmodium falciparum.

XX Synthetic.

XX WO9405311-A.

XX 17-MAR-1994.

XX 27-AUG-1993; 93WO-AU00441.

XX 27-AUG-1992; 92AU-0004374.

XX (DEAK-) DEAKIN RES LTD.

XX Comis A, Fischer P, Tyler MI;

XX WPI; 1994-100849/12.

XX New retro, Inverso and retro-inverso peptide analogues - used as  
 PT immunogens for the prodn. of antibodies and in diagnosis, therapy  
 PT and prophylaxis of diseases

XX Example 6; Page 21; 87pp; English.

XX Synthetic peptide antigen analogues of native peptide antigens which  
 CC are either (1) retro modified, (2) Inverso modified or (3) retro-  
 CC Inverso modified with respect to the native antigen can be used to  
 CC raise antibodies which recognise the native antigen and hence can be  
 CC used in the treatment and/or prophylaxis of diseases and therapy of  
 CC disease states. They can be used in the production of vaccines and  
 CC in assays for antibodies specific for the native antigen. They can  
 CC have increased stability to degradation after administration and they  
 CC can be administered orally. This sequence is based on the  
 CC immunodominant epitope of the circumsporozoite coat protein of  
 CC Plasmodium falciparum. The retro-inverso version of this sequence  
 CC is described in AAR60483.

XX Sequence 12 AA;

Query Match 100.0%; Score 69; DB 15; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
 | | | | | | | | | | | |  
 Db 1 nanpnanpnanp 12

RESULT 4

AAR70038  
 ID AAR70038 standard; Protein; 12 AA.

XX AC AAR70038;

XX 25-OCT-1995 (first entry)

XX Plasmodium falciparum circumsporozoite protein central repeat.

Qy 1 NANPNANPNANP 12  
 | | | | | | | | | | | |  
 Db 1 nanpnanpnanp 12



KW Bacteriophage Fd; phage display; malaria; vaccine; therapeutic;  
 XX diagnostic.

OS Synthetic.

XX WO9503454-A.

XX PD 23-FEB-1995.

XX 19-AUG-1994; 94WO-GB01827.

XX 19-AUG-1993; 93GB-0017304.

XX (CAMP-) CAMBRIDGE BACTERIOPHAGE TECHNOLOGIES.

XX Appella E, Perham RN, Veronese F, Willis AE;

XX WPI; 1995-098754/13.

XX DR N-PSDB; AAQ83024.

XX Filamentous bacteriophage engineered to display T- and B-cell  
 PT epitopes and/or a peptide eliciting HIV neutralising antibodies  
 PT - useful in vaccines and as a therapeutic/diagnostic prod.

XX PS Disclosure; Page 32; 42pp; English.

XX The protein represents 12 amino acid units of the P. falciparum  
 CC circumsporozoite protein repeat region (AAR70034), which are  
 CC sufficient for the production of antibodies reactive against the  
 CC whole protein. Insertion of the coding sequence into the major  
 CC coat protein gene VIII of bacteriophage Fd in a region of the  
 CC protein known to be exposed on the virus surface results in  
 CC expression of the P. falciparum antigen on the phage surface.  
 CC Thus, the phage is useful in a vaccine composition against  
 CC malaria, or as a therapeutic/diagnostic product.

XX SO Sequence 12 AA;

Query Match 100.0%; Score 69; DB 16; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.00016;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12

Db 1 nanpnanpnanp 12

RESULT 5

ID AAW61556 standard; peptide; 12 AA.

XX AAW61556;

XX 19-OCT-1998 (first entry)

XX Malaria circumsporozoite protein repeat region.

XX T-cell; malaria; immunogenic; anti-malarial; prophylactic immunity.

OS Plasmodium falciparum.

XX WO9831382-A1.

XX PD 23-JUL-1998.

XX 21-JAN-1998; 98WO-US01527.

XX 21-JAN-1997; 97US-0033916.

XX (UYNV ) UNIV NEW YORK STATE.

XX Moreno A, Nardin E;

XX WPI; 1998-413810/35.

XX New immunogenic compositions for malaria - comprise malaria derived  
 PT peptide comprising universal T-cell epitope which elicits  
 PT anti-malarial T-cell response

XX PS Disclosure; Page 2; 38pp; English.

XX The malaria circumsporozoite protein repeat region is used in an  
 CC immunogenic composition along with the T-cell epitope derived from  
 CC malaria. The T-cell epitope elicits an anti-malarial T-cell response  
 CC in mammals of diverse genetic backgrounds. The composition can be used  
 CC as a vaccine to confer prophylactic or therapeutic immunity against  
 CC malaria. They may also be used to inhibit the propagation of a malarial  
 CC organism in a susceptible animal.

XX SO Sequence 12 AA;

Query Match 100.0%; Score 69; DB 19; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.00016;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12

Db 1 nanpnanpnanp 12

RESULT 6

ID AAW29731 standard; peptide; 12 AA.

XX AAW29731;

XX 26-OCT-1998 (first entry)

XX Malarial B-cell epitope.

XX Circumsporozoite protein; CS; antibody; polyoxime; T-cell epitope;  
 KW vaccine; malaria.

XX Plasmodium falciparum.

XX WO9830237-A1.

XX PD 16-JUL-1998.

XX 24-DEC-1997; 97WO-US24283.

XX 24-DEC-1996; 96US-0034506.

XX (UYNV ) UNIV NEW YORK STATE.

XX Nardin E, Nussenzweig RS, Rose K;

XX WPI; 1998-398801/34.

XX Immunogenic composition containing T cell epitope of malaria-derived  
 PT peptide - incorporated into a polyoxime and optionally B cell  
 PT epitope, used in vaccines to protect against malaria in subjects of  
 PT differing genetic background(s)

XX Claim 7; p25; 36pp; English.

XX The present sequence is a peptide comprising a malarial B-cell epitope,  
 CC which is contained in the COOH-terminal repeat region of the Plasmodium  
 CC falciparum circumsporozoite (CS) protein. It is this epitope that  
 CC stimulates the production of antibodies that specifically recognise  
 CC and bind to the malarial CS protein. This epitope can be incorporated  
 CC into a polyoxime with T-cell epitopes to produce a vaccine that is  
 CC protective against malaria in individuals of different genetic  
 CC backgrounds.

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 69; DB 19; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
| | | | | | | | | | | |  
DB 1 nanpnanpnanp 12

RESULT 7  
RAY41790  
ID AAY41790 standard; peptide; 12 AA.  
XX AC AAY41790;  
XX DT 08-DEC-1999 (first entry)  
XX DE Plasmodium falciparum circumsporozoite core peptide.  
XX KW Immunogenic composition; viral particle; hybrid protein; expression;  
KW immunogen; diagnosis; passive immunisation; vaccination; vaccine;  
KW immune response; epitopic site; pathogen; toxin.  
XX OS Plasmodium falciparum.  
XX PN US5965140-A.  
XX PD 12-OCT-1999.  
XX PF 10-AUG-1998; 98US-0132079.  
XX PR 19-SEP-1985; 85US-0777976.  
PR 04-DEC-1987; 87US-0128639.  
PR 26-OCT-1992; 92US-0966291.  
PR 23-SEP-1993; 93US-0126115.  
PR 09-DEC-1994; 94US-0352989.  
PR 12-SEP-1984; 84US-0650323.  
XX (CHIR ) CHIRON CORP.  
XX PA Barr PJ, Kuo G, Valenzuela PDT;  
XX PI WPI; 1999-579892/49.  
XX DR N-PSDB; AA225062.  
XX New hybrid protein comprising a particle forming protein and a  
XX polypeptide having epitopes, useful as a vaccine -  
XX Disclosure; Column 20; 13pp; English.  
XX The present invention describes a hybrid protein (H1) comprising at  
XX least a portion of a particle forming protein and one or more  
XX polypeptides having at least one epitope of interest such as an  
XX epitopic site of a pathogen or toxin. Also described is a method  
XX for inducing an immune response to one or more epitopes of interest  
XX comprising immunising a host with a particle comprising at least in  
XX part H1 where the epitopic site is immunologically reactive. The  
XX particles can be used as vaccines or for the production of antibodies  
XX which can be used for passive immunisation, in vitro diagnostics,  
XX cytology, histology or other applications. The particles can also be  
XX used in assays for detecting the presence of antibodies to the antigen  
XX of interest. The present sequence represents a peptide used in the  
XX exemplification of the present invention.

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 69; DB 20; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00016;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
| | | | | | | | | | | |  
DB 1 nanpnanpnanp 12

RESULT 8  
AAY29734  
ID AAY29734 standard; Protein; 12 AA.  
XX AC AAY29734;  
XX DT 08-NOV-1999 (first entry)  
XX DE Plasmodium falciparum polypeptide haptan.  
XX KW Human hepatitis B core protein; Hbc; modified; immunodominant;  
KW nucleocapsid protein; vaccine; T cell epitope.  
XX OS Plasmodium falciparum.  
XX PN WO9940934-A1.  
XX PD 19-AUG-1999.  
XX PF 11-FEB-1999; 99WO-US03055.  
XX PR 12-FEB-1998; 98US-0074537.  
XX PA (IMMU-) IMMUNE COMPLEX CORP.  
XX PI Birkett AJ;  
XX WPI; 1999-527340/44.  
XX Conjugate of hepatitis B core protein, modified to increase  
XX reactivity with haptan, used to raise antibodies against the haptan,  
XX e.g. in vaccines  
XX Example 3; Page 39; 128pp; English.  
XX The present invention describes a conjugate (A) comprising a  
XX strategically modified hepatitis B core (Hbc) protein (I) attached to  
XX a haptan, where (I) includes amino acids (aa) 10-140 of the wild type  
XX Hbc 183 aa sequence (given in AAY29674) and additionally has an insert  
XX (II) in the region corresponding to aa's 50-100, where the insert is  
XX of 1 to about 40 aa's and contains a chemically reactive aa residue  
XX linked to the haptan. A vaccine containing (A), optionally in the form  
XX of particles, is used to induce a protective antibody response against  
XX the pathogen from which the haptan is derived, in humans or other  
XX animals. These pathogens may be bacteria, viruses, rickettsia or  
XX protozoa. Insertion of (II) overcomes the low reactivity of aa side  
XX chains in native Hbc protein, increasing the reactivity with haptan  
XX and resulting in conjugates of improved immunogenicity. Modified Hbc  
XX can be derivatised in the form of particles by well-defined chemical  
XX methods, and is unlikely to cause immunological side-effects. AAY29675  
XX to AAY29735 represent polypeptide haptans used in an example from the  
XX present invention.

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 69; DB 20; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
| | | | | | | | | | | |  
DB 1 nanpnanpnanp 12

RESULT 9

AAW67599  
 ID AAW67599 standard; peptide; 12 AA.  
 XX  
 AC AAW67599;  
 XX  
 DT 02-MAR-1999 (first entry)  
 XX  
 DE Plasmodium falciparum T-cell activation peptide.  
 XX  
 KW Activated T helper cell; CD4+; cytotoxic T cell; CD8+; liposome; epitope;  
 KW peripheral blood lymphocyte; antigen-presenting cell; APC; virus; tumour;  
 KW bacterium; parasite; cytokine; vaccine; cancer; malaria; HIV; hepatitis;  
 KW tuberculosis.  
 XX  
 OS Synthetic.  
 OS Plasmodium falciparum.  
 XX  
 PN WO9850527-A1.  
 XX  
 PD 12-NOV-1998.  
 XX  
 PF 07-MAY-1998; 98WO-US09288.  
 XX  
 PR 08-MAY-1997; 97US-0045949.  
 XX  
 PA (BIOM-) BIOMIRA INC.  
 XX  
 PI Agrawal B, Krantz MJ, Longenecker BM, Reddish MA;  
 XX  
 DR WPI; 1999-034715/03.  
 XX  
 PT Method of activation of T cells - by exposure to antigen-presenting  
 PT cells loaded with antigen in liposome, used for, e.g. treating  
 PT cancer and microbial infections  
 XX  
 PS Disclosure; Page 24; 75pp; English.  
 XX  
 CC Peptides AAW67583-W67611 are used to produce activated T helper (CD4+)  
 CC and cytotoxic (CD8+) T-cells. The activated T cells are produced by  
 CC treating peripheral blood lymphocytes with liposome-encapsulated peptide  
 CC antigen to generate Ag-loaded antigen-presenting cells (APC), contacting  
 CC naive or anergic T-cells with these APC, and isolating the resulting  
 CC activated T-cells. The cells are specific for a particular antigen,  
 CC particularly one derived from a tumour, but also those from viruses,  
 CC bacteria and other parasites. It can also be used to identify antigens  
 CC and epitopes able to generate an Ag-specific T-cell response (by  
 CC assessing proliferation and cytokine release). Also the Ag-loaded APC  
 CC can be used as cellular vaccines for treating cancer (claimed) or other  
 CC diseases (e.g. malaria, human immune deficiency virus infection,  
 CC hepatitis, tuberculosis). The activated T-cells can be used to treat the  
 CC same conditions by adoptive T-cell transfer therapy.  
 XX  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 69; DB 20; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
 Db 1 nanpnanpnnp 12  
 |||||

RESULT 10  
 AAY70280  
 ID AAY70280 standard; peptide; 12 AA.  
 XX  
 AC AAY70280;  
 XX  
 DT 06-JUN-2000 (first entry)  
 XX  
 DE Plasmodium falciparum CSP antigenic epitope, p519.

XX  
 KW Recombinant protein; CDC/NIIIMALVAC-1; multivalent; malaria; vaccine;  
 KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;  
 KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;  
 KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;  
 KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;  
 KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;  
 KW Pf27; antiparasitic; prevention; anti-CDC/NIIIMALVAC-1 antibody.  
 XX  
 OS Plasmodium falciparum.  
 XX  
 PN WO200011179-A1.  
 XX  
 PD 02-MAR-2000.  
 XX  
 PF 19-AUG-1999; 99WO-US18869.  
 XX  
 PR 21-AUG-1998; 98US-0097703.  
 XX  
 PA (NAIM-) NAT INST IMMUNOLOGY.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Lal AA, Shi YP, Hasnain SE;  
 XX  
 DR WPI; 2000-237654/20.  
 XX  
 PT Novel recombinant protein as vaccine for treating malarial infection  
 PT comprises antigenic peptides obtained from different stages of  
 PT plasmodium falciparum life cycle -  
 XX  
 PS Claim 2; Page 16; 52pp; English.  
 XX  
 CC The present sequence is the antigenic epitope P519, derived from  
 CC circumsporozoite protein (CSP) of the sporozoite stage of Plasmodium  
 CC falciparum. It is used in the construction of recombinant protein  
 CC CDC/NIIIMALVAC-1, which is a multivalent, multistage malarial vaccine.  
 CC The recombinant protein comprises, melittin signal peptide, (His)<sub>6</sub> tag,  
 CC T-cell epitope from tetanus toxoid and 21 antigenic epitopes from  
 CC circumsporozoite protein (CSP), sporozoite surface protein-2 (SSP-2),  
 CC liver stage antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1),  
 CC MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding antigen-175  
 CC (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete specific  
 CC antigen, Pf27. These epitopes were obtained at different stages of the  
 CC life cycle of P. falciparum. CDC/NIIIMALVAC-1 vaccine has antiparasitic  
 CC activity and can be used for treatment and prevention of malarial  
 CC infections. Anti-CDC/NIIIMALVAC-1 antibodies can be used for detecting  
 CC P. falciparum in biological samples.  
 XX  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 69; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
 Db 1 nanpnanpnnp 12  
 |||||

RESULT 11  
 AAB73753  
 ID AAB73753 standard; peptide; 12 AA.  
 XX  
 AC AAB73753;  
 XX  
 DT 21-SEP-2001 (first entry)  
 XX  
 DE Plasmodium falciparum circumsporozoite peptide antigen.  
 XX  
 KW Circumsporozoite peptide antigen; malaria parasite; epitope;  
 KW monoclonal antibody; affinity chromatography; DEAE-Sephadex A50 resin;  
 KW diethylamine-ethyl group.

Mon Feb 4 15:23:48 2002

us-09-763-397a-4.rag

XX OS Plasmodium falciparum.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1  
 XX FT /note= "Optionally acetylated"  
 XX PN WO200146216-A2.  
 XX PD 28-JUN-2001.  
 XX PF 20-DEC-2000; 2000WO-BR00160.  
 XX PR 22-DEC-1999; 99BR-0006091.  
 XX PA (CNPQ-) CNPQ CONSELHO NACIONAL DESENVOLVIMENTO.  
 XX PI Nakale CR, Ianzar DA, Cilli EM, Rodrigues MM;  
 XX DR WPI; 2001-475830/51.  
 XX PT Use of an anion exchange resin (EPM-7) as solid support for peptide,  
 XX PT e.g. angiotensin II, synthesis and affinity chromatography.  
 XX PS Disclosure; Page 6; 17pp; English.  
 XX CC The invention relates to the use of the anion exchange resin DEAE-  
 XX CC Sephadex A50 as a solid support for solid phase peptide synthesis and  
 XX CC affinity chromatography. The present sequence represents a repetitive  
 XX CC Plasmodium falciparum circumsporozoite peptide epitope which was used to  
 XX CC raise monoclonal antibodies. The affinity between the antibodies and free  
 XX CC recombinant antigen was determined via competition affinity  
 XX CC chromatography using a DEAE-Sephadex A50 column comprising a resin-bound  
 XX CC form of the peptide epitope (AAB73754).  
 XX SQ Sequence 12 AA;

Query Match 100.0%; Score 69; DB 22; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
 Db | | | | | | | | | | | |  
 1 nanpnanpnanp 12

RESULT 12  
 AAB84362  
 ID AAB84362 standard; peptide; 12 AA.  
 XX AC AAB84362;  
 XX DT 22-AUG-2001 (first entry)  
 XX DE Amino acid sequence of MALL1, which is displayed on E2 core surface.  
 XX DE E2 core protein; 2-oxo acid dehydrogenase; multienzyme complex;  
 XX KW vaccine; MALL1.  
 XX OS Unidentified.  
 XX OS WO200142439-A1.  
 XX PN 14-JUN-2001.  
 XX PD 08-DEC-2000; 2000WO-GB04720.  
 XX PF 09-DEC-1999; 99GB-0029151.  
 XX PR (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.  
 XX PA Perham RN, Domingo GJ;  
 XX PI

XX WPI; 2001-381670/40.  
 XX Novel truncated E2 core protein of 2-oxo acid dehydrogenase multienzyme  
 XX complex, which assembles into a core structure of the complex, useful  
 XX PT in screening for polypeptides which bind target proteins of interest -  
 XX PS Disclosure; Page 27; 53pp; English.  
 XX CC The specification describes a truncated E2 core protein of a 2-oxo  
 XX CC acid dehydrogenase multienzyme complex, where the protein assembles into  
 XX CC a core structure of the complex. The E2 core protein has an ability to  
 XX CC display a variety of different numbers of peptides or polypeptides on  
 XX CC a single complex, and the ability to display a variety of different  
 XX CC peptides or polypeptides on the same complex. The E2 core protein is  
 XX CC useful in screening for peptides or polypeptides which bind target  
 XX CC proteins of interest or bound by target antibodies, or which have other  
 XX CC desirable properties, and for the elicitation of immune responses,  
 XX CC e.g. for vaccination. It is also useful for obtaining and/or raising  
 XX CC antibodies to the peptides or polypeptides, for evoking an immune  
 XX CC response, and in pharmaceutical compositions. The present sequence  
 XX CC represents a MALL1 peptide, which is displayed on the E2 core protein  
 XX CC surface.  
 XX SQ Sequence 12 AA;

Query Match 100.0%; Score 69; DB 22; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
 Db | | | | | | | | | | | |  
 1 nanpnanpnanp 12

RESULT 13  
 AAB49997  
 ID AAB49997 standard; Peptide; 12 AA.  
 XX AC AAB49997;  
 XX DT 20-MAR-2001 (first entry)  
 XX DE Tumour antigen MUC-1 epitope peptide #1.  
 XX DE Somatic transgene immunisation; immune response; haematopoietic cell;  
 XX KW immunity; disease treatment; disease prevention.  
 XX OS Unidentified.  
 XX OS WO200064488-A2.  
 XX PN 02-NOV-2000.  
 XX PD 27-APR-2000; 2000WO-US11372.  
 XX PF 27-APR-1999; 99US-0300959.  
 XX PR (ZANE/) ZANETTI M.  
 XX PA Zanetti M;  
 XX PI WPI; 2001-024628/03.  
 XX DR Immunization by administering ex vivo a transgene having a  
 XX PT hematopoietic cell-specific expression element operationally linked to  
 XX PT a polynucleotide encoding one or more heterologous epitopes to a  
 XX PT lymphoid cell -  
 XX PS Example 8; Page 70; 100pp; English.  
 XX CC The present invention provides a novel method of immunisation involving

CC the administration of a haematopoietic cell specific expression element  
CC to a lymphoid cell. The method is known as somatic transgene  
CC immunisation. It is particularly useful in the treatment and prevention  
CC of tumours, autoimmune diseases, allergy, metabolic disorders, endocrine  
CC disorders, mental disorders, pain, blood disorders and dental disorders.  
XX  
SQ Sequence 12 AA;

Query Match 100.0%; Score 69; DB 22; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
| | | | | | | | | |  
DB 1 nanpnanpnanp 12

RESULT 14  
AAP60804  
ID AAP60804 standard; peptide; 13 AA.

XX  
AC AAP60804;

XX DT 22-JUL-1991 (first entry)

XX DE Sequence of an immunodominant epitope of P.falciparum  
DE circumsporozoite (CS) protein.

XX KW Malaria vaccine.

XX OS Plasmodium falciparum.

XX PN WO8605790-A.

XX PD 09-OCT-1986.

XX PF 27-MAR-1986; 86WO-US00627.

XX PR 28-MAR-1985; 85US-0716960.

XX PA (UYNV-) NEW YORK UNIVERSITY.

XX PI Nussenzweig VN, Zavala FP;

XX DR WPI; 1986-278805/42.

XX PT Conjugate of immunogenic peptide - having amino acid sequence  
PT corresp. to immuno-dominant epitope of P.falciparum  
PT circumsporozoite protein

XX PS Claim 18; Page 25; 37pp; English.

XX CC The patentors claim a novel conjugate which comprises AAP60804 and a  
CC carrier protein. The conjugates are useful for providing protective  
CC immunity against malaria. The conjugates are effective in raising  
CC high titres of antibodies in vivo. These antibodies recognise  
CC sporozoites and neutralise sporozoite infectivity in vitro by a  
CC vigorous CSP reaction. The epitope of the P.falciparum CS protein is  
CC not strain-specific.  
XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 69; DB 7; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
| | | | | | | | | |  
DB 2 nanpnanpnanp 13

RESULT 15  
AAR49980  
ID AAR49980 standard; peptide; 13 AA.

XX AC AAR49980;

XX DT 10-OCT-1994 (first entry)

XX DE Synthetic peptide analogue of P. falciparum used as immunogen.

XX KW Peptide; retro; Inverso; retro-inverso; therapy; disease; vaccine;  
KW immunogen; antibody; antibodies; assay; treatment; prophylaxis;  
KW Plasmodium falciparum.

XX OS Synthetic.

XX PN WO9405311-A.

XX PD 17-MAR-1994.

XX PF 27-AUG-1993; 93WO-AU000441.

XX PR 27-AUG-1992; 92AU-0004374.

XX PA (DEAK-) DEAKIN RES LTD.

XX PI Comis A, Fischer P, Tyler MI;

XX DR WPI; 1994-100849/12.

XX PT New retro, Inverso and retro-inverso peptide analogues - used as  
PT immunogens for the prodn. of antibodies and in diagnosis, therapy  
PT and prophylaxis of diseases

XX PS Example 6; Page 21; 87pp; English.

XX CC Synthetic peptide antigen analogues of native peptide antigens which  
CC are either (1) retro modified, (2) Inverso modified or (3) retro-  
CC Inverso modified with respect to the native antigen can be used to  
CC raise antibodies which recognise the native antigen and hence can be  
CC used in the treatment and/or prophylaxis of diseases and therapy of  
CC disease states. They can be used in the production of vaccines and  
CC in assays for antibodies specific for the native antigen. They can  
CC have increased stability to degradation after administration and they  
CC can be administered orally. This sequence is based on the  
CC immunodominant epitope of the circumsporozoite coat protein of  
CC Plasmodium falciparum. The retro-inverso version of this peptide  
CC is given in AAR60482.  
XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 69; DB 15; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
| | | | | | | | | |  
DB 2 nanpnanpnanp 13

Search completed: January 29, 2002, 10:21:41  
Job time: 417 sec

us-09-763-397a-4.rag

Mon Feb 4 15:23:48 2002

\_\_\_\_\_

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:01 ; Search time 133.18 Seconds  
(without alignments)  
2.028 Million cell updates/sec

Title: US-09-763-397A-4  
Perfect score: 69  
Sequence: 1 NANPNANPNANP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/1aa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	69	100.0	12	4	US-09-248-588-105
2	69	100.0	12	4	US-08-909-551-4
3	69	100.0	13	4	US-08-909-551-3
4	69	100.0	17	1	US-07-987-286-8
5	69	100.0	17	2	US-08-614-626-8
6	69	100.0	24	6	5219987-4
7	69	100.0	66	2	US-08-455-625-35
8	69	100.0	66	4	US-08-455-685-35
9	69	100.0	66	4	US-08-060-988A-35
10	69	100.0	66	5	PCT-US94-05142-35
11	69	100.0	68	1	US-08-143-365A-13
12	69	100.0	126	1	US-08-268-348A-12
13	69	100.0	126	1	US-08-395-602A-4
14	69	100.0	126	2	US-08-021-625D-4
15	69	100.0	133	1	US-08-268-348A-8
16	69	100.0	133	1	US-08-268-348A-10
17	69	100.0	160	6	5219987-7
18	69	100.0	163	5	PCT-US93-08435-2
19	69	100.0	164	5	PCT-US93-08435-1
20	69	100.0	412	1	US-08-313-288B-18
21	69	100.0	423	2	US-08-760-797A-1
22	69	100.0	424	2	US-08-760-797A-3
23	69	100.0	424	4	US-08-932-929B-1
24	69	100.0	424	4	US-08-932-929B-3
25	60	87.0	46	1	US-08-395-602A-5
26	60	87.0	46	2	US-08-021-625D-5
27	59	85.5	12	2	US-08-747-137-41

28	54	78.3	28	3	US-08-822-324-23	Sequence 23, Appl
29	54	78.3	1687	2	US-08-570-311-29	Sequence 29, Appl
30	52	75.4	9	2	US-08-318-856A-13	Sequence 13, Appl
31	49	71.0	659	4	US-08-894-818B-1	Sequence 1, Appl
32	49	71.0	659	4	US-08-894-818B-5	Sequence 5, Appl
33	49	71.0	1087	2	US-08-570-311-8	Sequence 8, Appl
34	49	71.0	1087	2	US-08-353-485-8	Sequence 27, Appl
35	49	71.0	1358	2	US-08-570-311-27	Sequence 8, Appl
36	49	71.0	1704	3	US-08-336-308A-10	Sequence 10, Appl
37	49	71.0	1704	3	US-08-822-324-6	Sequence 6, Appl
38	49	71.0	1704	4	US-09-490-931-10	Sequence 10, Appl
39	49	71.0	1732	2	US-08-570-311-10	Sequence 10, Appl
40	49	71.0	1732	2	US-08-353-485-10	Sequence 10, Appl
41	49	71.0	2628	2	US-08-570-311-14	Sequence 14, Appl
42	46	66.7	8	2	US-08-747-137-40	Sequence 40, Appl
43	46	66.7	8	2	US-08-318-856A-12	Sequence 12, Appl
44	46	66.7	16	1	US-08-346-849-50	Sequence 50, Appl
45	46	66.7	16	2	US-08-293-284A-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1  
US-09-248-588-105  
; Sequence 105, Application US/09248588  
; Patent No. 6231864  
; GENERAL INFORMATION:  
; APPLICANT: Birkett, Ashley J.  
; TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and  
; FILE OF INVENTION: their Derivatives  
; FILE REFERENCE: SYN-101 4564/69529  
; CURRENT APPLICATION NUMBER: US/09/248,588  
; CURRENT FILING DATE: 1999-02-11  
; EARLIER APPLICATION NUMBER: 60/074537  
; EARLIER FILING DATE: 1998-02-12  
; NUMBER OF SEQ ID NOS: 113  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 105  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Cytochrome  
; OTHER INFORMATION: P-450 fragment  
; PUBLICATION INFORMATION:  
; JOURNAL: Science  
; VOLUME: 228  
; PAGES: 1436-1440  
; DATE: 1985  
US-09-248-588-105

Query Match 100.0%; Score 69; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 8.3e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
DB 1 NANPNANPNANP 12

RESULT 2  
US-08-909-551-4  
; Sequence 4, Application US/08909551  
; Patent No. 6261569  
; GENERAL INFORMATION:  
; APPLICANT: COMIS, ALFIO  
; APPLICANT: FISCHER, PETER  
; TITLE OF INVENTION: Retro-, Inverso-, and Retro-Inverso  
; TITLE OF INVENTION: Synthetic Peptide Analogues  
; NUMBER OF SEQUENCES: 16

us-09-763-397a-4.ra1

Mon Feb 4 15:23:48 2002

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., P.O. Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: United States of America
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Vers. #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/909,551
FILING DATE: 12-AUG-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/387,932
FILING DATE: 27-FEB-1995
APPLICATION NUMBER: PCT/AU93/00441
FILING DATE: 27-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU 4374
FILING DATE: 27-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GHC20USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
DEVELOPMENTAL STAGE: SPOROZOITE
PUBLICATION INFORMATION:
DOCUMENT NUMBER: AU PL4374
FILING DATE: 27-AUG-1992
US-08-909-551-4

Query Match 100.0%; Score 69; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.3e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12
Db 1 NANPNANPNANP 12

RESULT 3
US-08-909-551-3
; Sequence 3, Application US/080909551
; Patent No. 6261569
; GENERAL INFORMATION:
; APPLICANT: COMIS, ALFIO
; APPLICANT: FISCHER, PETER
; TITLE OF INVENTION: Retro-, Inverso-, and Retro-Inverso
; TITLE OF INVENTION: Synthetic Peptide Analogues
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457

CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., P.O. Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: United States of America
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Vers. #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/909,551
FILING DATE: 12-AUG-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/387,932
FILING DATE: 27-FEB-1995
APPLICATION NUMBER: PCT/AU93/00441
FILING DATE: 27-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU 4374
FILING DATE: 27-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GHC20USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
DEVELOPMENTAL STAGE: SPOROZOITE
PUBLICATION INFORMATION:
DOCUMENT NUMBER: AU PL4374
FILING DATE: 27-AUG-1992
US-08-909-551-3

Query Match 100.0%; Score 69; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12
Db 2 NANPNANPNANP 13

RESULT 4
US-07-987-286-8
; Sequence 8, Application 07/987286
; Patent No. 5500366
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, GREGORY J.
; APPLICANT: GECZY, ANDREW F.
; TITLE OF INVENTION: T-CELL EPITOPES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; STREET: Spring House Corporate Cntr., P.O. Box 457
```



ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K STREET, N.W.  
CITY: WASHINGTON, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 07/987,286  
FILING DATE: 16-MAR-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU91/00429  
FILING DATE: 17-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PK2361  
FILING DATE: 18-SEP-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, STEPHEN A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/175/CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-987-286-8

Query Match 100.0%; Score 69; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.00012;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 4 NANPNANPNANP 15

RESULT 5  
US-08-614-626-8  
Sequence 8, Application US/08614626  
Patent No. 5928644  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, GREGORY J  
APPLICANT: GECZY, ANDREW F  
TITLE OF INVENTION: T-CELL EPITOPES  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K STREET, N.W.  
CITY: WASHINGTON, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/614,626  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/987,286  
FILING DATE: 16-MAR-1993

APPLICATION NUMBER: PCT/AU91/00429  
FILING DATE: 17-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PK2361  
FILING DATE: 18-SEP-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, STEPHEN A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/175/CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-614-626-8

Query Match 100.0%; Score 69; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.00012;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 4 NANPNANPNANP 15

RESULT 6  
5219987-4  
Patent No. 5219987  
APPLICANT: VERDINI, ANTONIO S.; PESSI, ANTONELLO; BONELLI,  
FABIO  
TITLE OF INVENTION: SEQUENTIAL POLYPEPTIDES ENDOWED WITH  
IMMUNOLOGICAL ACTIVITY  
NUMBER OF SEQUENCES: 7  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/227,364  
FILING DATE: 02-AUG-1988  
SEQ ID NO: 4  
LENGTH: 24  
5219987-4

Query Match 100.0%; Score 69; DB 6; Length 24;  
Best Local Similarity 100.0%; Pred. No. 0.00017;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 3 NANPNANPNANP 14

RESULT 7  
US-08-455-625-35  
Sequence 35, Application US/08455625  
Patent No. 5932218  
GENERAL INFORMATION:  
APPLICANT: Berzofsky, Jay A.  
APPLICANT: Ahlers, Jeffrey D.  
APPLICANT: Pendleton, C. D.  
APPLICANT: Nara, Peter  
APPLICANT: Shirai, Mutsunori  
TITLE OF INVENTION: COMPOSITE SYNTHETIC PEPTIDE CONSTRUCT  
ELICITING NEUTRALIZING ANTIBODIES AND CYTOTOXIC T  
CELLS  
TITLE OF INVENTION: LYMPHOCYTES AGAINST HIV  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: P.O. Box 747

Mon Feb 4 15:23:48 2002

CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,625  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/060,988  
FILING DATE: 14-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30330  
REFERENCE/DOCKET NUMBER: 1173-434P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 66 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..66  
OTHER INFORMATION: /label= peptide  
OTHER INFORMATION: /note= "peptide from P. falciparum CS antigen"  
US-08-455-625-35

Query Match 100.0%; Score 69; DB 2; Length 66;  
Best Local Similarity 100.0%; Pred. No. 0.00052;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 22 NANPNANPNANP 33

RESULT 8  
US-08-455-685-35  
Sequence 35, Application US/08455685  
Patent No. 6214347  
GENERAL INFORMATION:  
APPLICANT: Berzofsky, Jay A.  
APPLICANT: Ahlers, Jeffrey D.  
APPLICANT: Pendleton, C. David  
APPLICANT: Nara, Peter  
APPLICANT: Shirai, Mutsunori  
TITLE OF INVENTION: MULTIDETERMINANT PEPTIDES THAT ELICIT  
TITLE OF INVENTION: HELPER T-LYMPHOCYTE, CYTOTOXIC T LYMPHOCYTE AND  
TITLE OF INVENTION: NEUTRALIZING ANTIBODY RESPONSES AGAINST HIV-1  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,685  
FILING DATE: 31-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/060,988  
FILING DATE: 14-MAY-1993  
APPLICATION NUMBER: 07/847,311  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/751,998  
FILING DATE: 29-AUG-1991  
APPLICATION NUMBER: 07/148,692  
FILING DATE: 26-JAN-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Beattie, Ingrid A.  
REGISTRATION NUMBER: P-42,306  
REFERENCE/DOCKET NUMBER: 08830/022003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 66 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-455-685-35

Query Match 100.0%; Score 69; DB 4; Length 66;  
Best Local Similarity 100.0%; Pred. No. 0.00052;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 22 NANPNANPNANP 33

RESULT 9  
US-08-060-988A-35  
Sequence 35, Application US/08060988A  
Patent No. 6294322  
GENERAL INFORMATION:  
APPLICANT: Berzofsky, Jay A.  
APPLICANT: Ahlers, Jeffrey D.  
APPLICANT: Pendleton, C. David  
APPLICANT: Nara, Peter  
APPLICANT: Shirai, Mutsunori  
TITLE OF INVENTION: MULTIDETERMINANT PEPTIDES  
TITLE OF INVENTION: THAT ELICIT  
TITLE OF INVENTION: HELPER T-LYMPHOCYTE, CYTOTOXIC T LYMPHOCYTE AND  
TITLE OF INVENTION: NEUTRALIZING ANTIBODY RESPONSES AGAINST HIV-1  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/060,988A  
FILING DATE: 14-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/847,311  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/751,998  
FILING DATE: 29-AUG-1991

APPLICATION NUMBER: 07/148,692  
FILING DATE: 26-JAN-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Beattie, Ingrid A.  
REGISTRATION NUMBER: P-42,306  
REFERENCE/DOCKET NUMBER: 08830/022001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 66 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-060-988A-35

Query Match 100.0%; Score 69; DB 4; Length 66;  
Best Local Similarity 100.0%; Pred. No. 0.00052;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
Db 22 NANPNANPNANP 33

RESULT 10  
PCT-US94-05142-35  
; Sequence 35, Application PC/TUS9405142  
; GENERAL INFORMATION:  
; APPLICANT: COMPOSITE SYNTHETIC PEPTIDE CONSTRUCT  
; TITLE OF INVENTION: ELICITING NEUTRALIZING ANTIBODIES AND CYTOTOXIC T  
; TITLE OF INVENTION: LYMPHOCYTES AGAINST HIV  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: P.O. Box 747  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05142  
; FILING DATE: 13-MAY-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/060,988  
; FILING DATE: 14-MAY-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svensson, Leonard R.  
; REGISTRATION NUMBER: 30330  
; REFERENCE/DOCKET NUMBER: 1173-434P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-205-8000  
; TELEFAX: 703-205-8050  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 66 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..66

OTHER INFORMATION: /label= peptide  
OTHER INFORMATION: /note= "peptide from P. falciparum CS antigen"  
PCT-US94-05142-35

Query Match 100.0%; Score 69; DB 5; Length 66;  
Best Local Similarity 100.0%; Pred. No. 0.00052;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
Db 22 NANPNANPNANP 33

RESULT 11  
US-08-143-365A-13  
; Sequence 13, Application US/08143365A  
; Patent No. 5726292  
; GENERAL INFORMATION:  
; APPLICANT: Lowell, George H  
; TITLE OF INVENTION: IMMUNO-POTENTIATING SYSTEMS FOR  
; TITLE OF INVENTION: PREPARATION OF IMMUNOGENIC MATERIALS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hendricks and Assoc.  
; STREET: 9669 A Main Street  
; CITY: Fairfax  
; STATE: VA  
; COUNTRY: US  
; ZIP: 22031  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143.365A  
; FILING DATE: 29-OCT-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hendricks, Glenna M  
; REGISTRATION NUMBER: 32,535  
; REFERENCE/DOCKET NUMBER: Lowell-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 425-4250  
; TELEFAX: (703) 425-2767  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 68 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-143-365A-13

Query Match 100.0%; Score 69; DB 1; Length 68;  
Best Local Similarity 100.0%; Pred. No. 0.00053;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
Db 1 NANPNANPNANP 12

RESULT 12  
US-08-268-348A-12  
; Sequence 12, Application US/08268348A  
; Patent No. 5750374  
; GENERAL INFORMATION:

Mon Feb 4 15:23:48 2002

APPLICANT: Dobelli, Heinz  
APPLICANT: Draeger, Nicholas  
APPLICANT: Trotman, Gerda H  
APPLICANT: Jakob, Peter  
APPLICANT: Stuber, Dietrich  
TITLE OF INVENTION: Process for Producing Hydrophobic  
TITLE OF INVENTION: Polypeptides and Proteins, and Fusion Proteins for Use in  
TITLE OF INVENTION: Producing Same  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann-La Roche Inc.  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/268,348A  
FILING DATE: 29-JUN-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 93110755.1  
FILING DATE: 06-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Parise, John P.  
REGISTRATION NUMBER: 34,403  
REFERENCE/DOCKET NUMBER: 4105/157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-6326  
TELEFAX: (201) 235-3500  
INFORMATION FOR SEQ ID NO: 12:  
LENGTH: 126 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-268-348A-12

Query Match 100.0%; Score 69; DB 1; Length 126;  
Best Local Similarity 100.0%; Pred. NO. 0.001;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
DB 13 NANPNANPNANP 24

RESULT 13  
US-08-395-602A-4  
Sequence 4, Application US/08395602A  
Patent No. 5766899  
GENERAL INFORMATION:  
APPLICANT: Kuo, M. Tien  
APPLICANT: Ding, Zhi-Ming  
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into  
TITLE OF INVENTION: Liver Cells  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/395,602A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: UTSC:410/HYL  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 126 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-395-602A-4

Query Match 100.0%; Score 69; DB 1; Length 126;  
Best Local Similarity 100.0%; Pred. NO. 0.001;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
DB 25 NANPNANPNANP 36

RESULT 14  
US-08-021-625D-4  
Sequence 4, Application US/08021625D  
Patent No. 5976851  
GENERAL INFORMATION:  
APPLICANT: Kuo, M. Tien  
APPLICANT: Ding, Zhi-Ming  
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into  
TITLE OF INVENTION: Liver Cells  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/021,625D  
FILING DATE: 16-FEB-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: UTSC:410/HYL  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 126 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-021-625D-4

Query Match 100.0%; Score 69; DB 2; Length 126;  
Best Local Similarity 100.0%; Pred. No. 0.001;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
      |  
Db 25 NANPNANPNANP 36  
      |

RESULT 15  
US-08-268-348A-8  
; Sequence 8, Application US/08268348A  
; Patent No. 5750374  
; GENERAL INFORMATION:  
; APPLICANT: Dobeli, Heinz  
; APPLICANT: Draeger, Nicholas  
; APPLICANT: Trotman, Gerda H  
; APPLICANT: Jakob, Peter  
; APPLICANT: Stuber, Dietrich  
; TITLE OF INVENTION: Process for Producing Hydrophobic  
; TITLE OF INVENTION: polypeptides and Proteins, and Fusion Proteins for Use in  
; TITLE OF INVENTION: Producing Same  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann-La Roche Inc.  
; STREET: 340 Kingsland Street  
; CITY: Nutley  
; STATE: New Jersey  
; COUNTRY: U.S.A.  
; ZIP: 07110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/268,348A  
; FILING DATE: 29-JUN-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 93110755.1  
; FILING DATE: 06-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parlise, John P.  
; REGISTRATION NUMBER: 34,403  
; REFERENCE/DOCKET NUMBER: 4105/157  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (201) 235-6326  
; TELEFAX: (201) 235-3500  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 133 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-268-348A-8

Query Match 100.0%; Score 69; DB 1; Length 133;  
Best Local Similarity 100.0%; Pred. No. 0.001;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
      |  
Db 13 NANPNANPNANP 24  
      |

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:34 ; Search time 144.96 seconds  
(without alignments)  
6.306 Million cell updates/sec

Title: US-09-763-397A-4

Perfect score: 69

Sequence: 1 NANPNANPNANP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	69	100.0	388	2	A39756	circumsporozoite p
2	69	100.0	405	2	S05428	circumsporozoite p
3	69	100.0	412	1	OZQAF	circumsporozoite p
4	69	100.0	424	2	A54533	circumsporozoite p
5	69	100.0	442	2	A54529	circumsporozoite p
6	54	78.3	387	2	A86322	hypothetical prote
7	54	78.3	1526	2	S49763	hypothetical prote
8	51	73.9	171	2	A29795	gingipain R (EC 3.
9	51	73.9	189	2	B29795	circumsporozoite p
10	50	72.5	1171	2	T13065	circumsporozoite p
11	49	71.0	1528	2	D85912	PIP82 protein - fr
12	49	71.0	1704	2	A55426	hypothetical prote
13	49	71.0	1732	2	T30836	gingipain R (EC 3.
14	49	71.0	2628	2	T28651	lysine-specific cy
15	48	69.6	790	2	T12203	hemagglutinin A -
16	48	69.6	967	2	H86334	transcription fact
17	46	66.7	182	2	S73046	hypothetical prote
18	46	66.7	953	2	A86351	penicillin binding
19	45	65.2	169	2	T41886	hypothetical prote
20	45	65.2	501	2	T48336	hypothetical prote
21	44	63.8	346	2	T01123	hypothetical prote
22	44	63.8	501	2	B81048	hypothetical prote
23	44	63.8	517	2	G81060	tsap protein, prob
24	44	63.8	548	2	A81070	tsap protein, prob
25	43	62.3	356	2	A96826	TK14.10 [imported
26	43	62.3	402	2	S31196	hypothetical prote
27	43	62.3	441	2	T12011	cellulase (EC 3.2.
28	43	62.3	898	2	T49492	hypothetical prote
29	43	62.3	1349	2	E85524	probable beta-barr

30	43	62.3	2150	2	S71629	sensory transducti
31	42	60.9	115	2	T20983	hypothetical prote
32	42	60.9	263	2	B84778	hypothetical prote
33	42	60.9	474	2	T21771	hypothetical prote
34	42	60.9	818	1	JC4397	peroxinectin precu
35	41.5	60.1	306	2	H71925	cag island protein
36	41	59.4	162	1	YA2051	antigen 5.1 precu
37	41	59.4	162	2	A26769	antigen 5.1 precu
38	41	59.4	162	2	A25780	blood-stage antige
39	41	59.4	252	2	T45737	hypothetical prote
40	41	59.4	289	2	T49024	hypothetical prote
41	41	59.4	442	2	G84778	probable bzif tran
42	41	59.4	1033	2	T42701	hypothetical prote
43	41	59.4	1314	1	TNBYR6	transcription regu
44	40	58.0	64	2	PC1237	peptidylprolyl iso
45	40	58.0	169	2	S48760	SNF1 protein - ye

#### ALIGNMENTS

```

RESULT 1
A39756
C:Species: Plasmodium reichenowi
C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999
C:Accession: A39756
R:Lal, A.A.: Goldman, I.F.
J. Biol. Chem. 266, 6686-6689, 1991
A:Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malar
A:Reference number: A39756; MUID:91201303
A:Accession: A39756
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <LAL>
A:Cross-references: GB:M60972; NID:g160228; PIDN:AAA29561.1; PID:g160229
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:312-366/Domain: thrombospondin type 1 repeat homology <THRA>

Query Match      100.0%; Score 69; DB 2; Length 388;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12
Db 196 NANPNANPNANP 207

RESULT 2
S05428
C:Species: Plasmodium falciparum
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000
C:Accession: S05428; A45527; I60657
R:Campbell, J.R.
Nucleic Acids Res. 17, 5854, 1989
A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate
A:Reference number: S05428
A:Accession: S05428
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-405 <CAM>
A:Cross-references: EMBL:X15363
A:Caspar, P.; Gentz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.
Mol. Biochem. Parasitol. 35, 185-190, 1989
A:Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate
A:Reference number: A45527; MUID:89364998
A:Accession: A45527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <CAS>
A:Cross-references: GB:M22982; GB:J04650; NID:g160168; PIDN:AAA29527.1; PID:g160169

```

R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.  
 Mol. Biochem. Parasitol. 37, 275-280, 1989  
 A:Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell epitopes  
 A:Reference number: A60657; MUID:90114334  
 A:Accession: 160657  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 319-336,354-373 <LOC>  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 F:329-383/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 69; DB 2; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 0.0021;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 |||||  
 DB 129 NANPNANPNANP 140

RESULT 3  
 ORZQAF  
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate IMW22)  
 C:Species: Plasmodium falciparum  
 C:Date: 15-Nov-1984 #sequence\_revision 15-Nov-1984 #text\_change 09-Jun-2000  
 C:Accession: A03388  
 R:Dame, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W.T.  
 Science 225, 593-599, 1984  
 A:Title: Structure of the gene encoding the immunodominant surface antigen on the sporozoite  
 A:Reference number: A03388; MUID:84250215  
 A:Accession: A03388  
 A:Molecule type: DNA  
 A:Residues: 1-412 <DAM>  
 A:Cross-references: GB:K02194; NID:g160160; PIDN:AAA29524.1; PID:g160161  
 A:Experimental source: clone 768  
 C:Comment: Residues 1-16 are the probable signal sequence.  
 C:Comment: There are 41 copies of a 4-residue repeating unit in the middle domain of the C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 F:336-390/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 69; DB 1; Length 412;  
 Best Local Similarity 100.0%; Pred. No. 0.0021;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 |||||  
 DB 148 NANPNANPNANP 159

RESULT 4  
 A54533  
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, Thailand)  
 C:Species: Plasmodium falciparum  
 C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jun-2000  
 C:Accession: A54533  
 R:del Portillo, H.A.; Nussenzeig, R.S.; Enea, V.  
 Mol. Biochem. Parasitol. 24, 289-294, 1987  
 A:Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.  
 A:Reference number: A54533; MUID:87315205  
 A:Accession: A54533  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-424 <DEL>  
 A:Cross-references: GB:M19752; NID:g160216; PIDN:AAA29555.1; PID:g160217  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 F:348-402/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 69; DB 2; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 0.0022;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 |||||  
 DB 148 NANPNANPNANP 159

RESULT 5  
 A54529  
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain Wellcome)  
 C:Species: Plasmodium falciparum  
 C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jun-2000  
 C:Accession: A54529  
 R:Lockyer, M.J.; Schwarz, R.T.  
 Mol. Biochem. Parasitol. 22, 101-108, 1987  
 A:Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum  
 A:Reference number: A54529; MUID:87115616  
 A:Accession: A54529  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-442 <LOC>  
 A:Cross-references: GB:M15505; NID:g160214; PIDN:AAA29554.1; PID:g160215  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 C:Keywords: tandem repeat  
 F:366-420/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 69; DB 2; Length 442;  
 Best Local Similarity 100.0%; Pred. No. 0.0023;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 |||||  
 DB 150 NANPNANPNANP 161

RESULT 6  
 A86322  
 hypothetical protein AAF27098.1 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: A86322  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Hultzer, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: A86322  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-387 <STO>  
 A:Cross-references: GB:AE005172; NID:g6730703; PIDN:AAF27098.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 78.3%; Score 54; DB 2; Length 387;  
 Best Local Similarity 75.0%; Pred. No. 0.35;  
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 |||||  
 DB 84 NANPNANPNANP 95

RESULT 7  
 S49763  
 gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis (fragment)



C:Species: Porphyromonas gingivalis  
 C:Date: 05-Mar-1995 #sequence\_revision 12-May-1995 #text\_change 31-Mar-1997  
 C:Accession: S49763  
 R:Aduse-Opoku, J.; Mullr, J.; Stanley, J.M.; Rangarajan, M.; Curtis, M.A.  
 Submitted to the EMBL Data Library, November 1994  
 A:Description: Cloning, sequence analysis and expression in Escherichia coli of prp1 of  
 A:Reference number: S49763  
 A:Accession: S49763  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1526 <ADU>  
 A:Cross-references: EMBL:X82680  
 C:Genetics:  
 A:Gene: prp1  
 C:Keywords: cysteine proteinase; hydrolase

Query Match 78.3%; Score 54; DB 2; Length 1526;  
 Best Local Similarity 75.0%; Pred. No. 1.5;  
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 | | | | | | | | | |  
 Db 951 NPNPNPNPNPNP 962

RESULT 8  
 A29795  
 C:Species: Plasmodium falciparum - malaria parasite (Plasmodium falciparum) (strain B11) (fragment)  
 C:Date: 28-Aug-1989 #sequence\_revision 28-Aug-1989 #text\_change 09-Jun-2000  
 C:Accession: A29795  
 R:de la Cruz, V.F.; Lal, A.A.; McCutchan, T.F.  
 J. Biol. Chem. 262, 11935-11939, 1987  
 A:Title: Sequence variation in putative functional domains of the circumsporozoite prote  
 A:Reference number: A92609; MUID:87308186  
 A:Accession: A29795  
 A:Molecule type: DNA  
 A:Residues: 1-171 <DEFLA>  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

Query Match 73.9%; Score 51; DB 2; Length 171;  
 Best Local Similarity 90.0%; Pred. No. 0.41;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ANPNANPNAN 11  
 | | | | | | | | | |  
 Db 117 ANPNANPNKN 126

RESULT 9  
 B29795  
 C:Species: Plasmodium falciparum - malaria parasite (Plasmodium falciparum) (strain LB5) (fragment)  
 C:Date: 18-Oct-1989 #sequence\_revision 21-Jul-1995 #text\_change 09-Jun-2000  
 C:Accession: B29795; A60657  
 R:de la Cruz, V.F.; Lal, A.A.; McCutchan, T.F.  
 J. Biol. Chem. 262, 11935-11939, 1987  
 A:Title: Sequence variation in putative functional domains of the circumsporozoite prote  
 A:Reference number: A92609; MUID:87308186  
 A:Accession: B29795  
 A:Molecule type: DNA  
 A:Residues: 1-161 <DEFLA>  
 A:Cross-references: GB:M17802; GB:M17803; GB:M17806  
 R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.  
 Mol. Biochem. Parasitol. 37, 275-280, 1989  
 A:Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell ep  
 A:Reference number: A60657; MUID:90114334  
 A:Accession: A60657  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 152-189 <LOC>

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

Query Match 73.9%; Score 51; DB 2; Length 189;  
 Best Local Similarity 90.0%; Pred. No. 0.46;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ANPNANPNAN 11  
 | | | | | | | | | |  
 Db 107 ANPNANPNKN 116

RESULT 10  
 T13065  
 C:Species: Drosophila melanogaster  
 C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
 C:Accession: T13065  
 R:Surl, V.; Qian, Z.; Hall, J.C.; Rosbash, M.  
 Neuron 21, 225-234, 1998  
 A:Title: Evidence that TIM light response is relevant to light-induced phase shifts 1  
 A:Reference number: Z17597; MUID:98561331  
 A:Accession: T13065  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1171 <SUR>  
 A:Cross-references: EMBL:AF067153; NID:93171244; PID:93171245; PIDN:AA018395.1  
 A:Experimental source: strain Canton-S, photoreceptor  
 C:Genetics:  
 A:Cross-references: FlyBase:FBgn0024943

Query Match 72.5%; Score 50; DB 2; Length 1171;  
 Best Local Similarity 66.7%; Pred. No. 4.6;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 | | | | | | | | | |  
 Db 1113 NPNPNPNPNPNP 1124

RESULT 11  
 D85912  
 C:Species: Escherichia coli (strain 0157:H7)  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
 C:Accession: D85912  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
 iller, L.; Grobeck, E.J.; Davis, N.W.; Llim, A.; Diallanta, E.; Potamousis, K.; Apoda  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206531  
 A:Accession: D85912  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1528 <SUR>  
 A:Cross-references: GB:AE005174; NID:912517083; PIDN:AA057760.1; GSPDB:GN00145; UNCP:  
 A:Experimental source: strain 0157:H7, substrain EDJ933  
 C:Genetics:  
 A:Gene: yp1a

Query Match 71.0%; Score 49; DB 2; Length 1528;  
 Best Local Similarity 66.7%; Pred. No. 8.7;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 | | | | | | | | | |  
 Db 1185 NPDNPNPNPNP 1196

RESULT 12  
 A55426

gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis  
N:Alternate names: 50K high molecular mass arginine-specific cysteine proteinase; HGP; R  
C:Species: Porphyromonas gingivalis  
C>Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 08-Oct-1999  
C:Accession: A55426; D53113  
R:Pavloff, N.; Potempa, J.; Pike, R.N.; Prochazka, V.; Kiefer, M.C.; Travis, J.; Barr, F.  
J. Biol. Chem. 270, 1007-1010, 1995  
A:Title: Molecular cloning and structural characterization of the Arg-gingipain proteol  
A:Reference number: A55426; MUID:9518080  
A:Accession: A55426  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1704 <PAV>  
A:Cross-references: GB:U15282; NID:9557067; PIDN:AAA69539.1; PID:9557068  
R:Pike, R.; McGraw, W.; Potempa, J.; Travis, J.  
J. Biol. Chem. 269, 406-411, 1994  
A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isol  
A:Reference number: A53113; MUID:94103245  
A:Accession: D53113  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 228-249 <PIK>  
A:Experimental source: H66  
A:Note: sequence extracted from NCBI backbone (NCBIP:141694)  
C:Keywords: cysteine proteinase; hydrolase

Query Match 71.0%; Score 49; DB 2; Length 1704;  
Best Local Similarity 66.7%; Pred. No. 9.8;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 947 NGTPNPMPNPMP 958

RESULT 13  
T30836  
Lysine-specific cysteine proteinase porphyain (EC 3.4.22.-) - Porphyromonas gingivalis  
N:Alternate names: Lysine-specific cysteine proteinase 1, 60K  
C:Species: Porphyromonas gingivalis  
C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 17-Nov-2000  
C:Accession: T30836; T30837; T30526; A53113  
R:Barokcy-Gallagher, G.A.; Han, N.; Patil, J.M.; Whitlock, J.; Proguiske-Fox, A.; Lantz,  
J. Bacteriol. 178, 2734-2741, 1996  
A:Title: Analysis of the prtP gene encoding porphyain, a cysteine proteinase of Porphy  
A:Reference number: Z20895; MUID:96213011  
A:Accession: T30836  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1732 <BAR>  
A:Cross-references: EMBL:U42210; NID:g1314325; PID:g1314326; PIDN:AAB06565.1  
R:Slaski, N.; Cleal, S.M.; Reynolds, E.C.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z20896  
A:Accession: T30837  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-795; 'I', 797-1389; 'N', 1391-1478; 'Y', 1480-1732 <SLA>  
A:Cross-references: EMBL:U75366; NID:g2182811; PID:g2182812; PIDN:AAB60809.1  
R:Lewis, J.P.; Macrina, F.L.  
Infect. Immun. 66, 3035-3042, 1998  
A:Title: IS195, an insertion sequence-like element associated with protease genes in Por  
A:Reference number: Z20844; MUID:98298016  
A:Accession: T30526  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1350; 'N', 1352-1363; 'Y', 1365-1447; 'H', 1449-1732 <LEW>  
A:Cross-references: EMBL:AF017059; NID:g2738802; PID:g2738803; PIDN:AAC26523.1  
R:Pike, R.; McGraw, W.; Potempa, J.; Travis, J.  
J. Biol. Chem. 269, 406-411, 1994  
A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isol  
A:Reference number: A53113; MUID:94103245

A:Accession: A53113  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 229-249 <PIK>  
A:Experimental source: H66  
A:Note: sequence extracted from NCBI backbone (NCBIP:141690)  
C:Genetics:  
A:Gene: prtP  
C:Keywords: cysteine proteinase; hydrolase

Query Match 71.0%; Score 49; DB 2; Length 1732;  
Best Local Similarity 66.7%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 968 NGTPNPMPNPMP 979

RESULT 14  
T28651  
hemagglutinin A - Porphyromonas gingivalis  
C:Species: Porphyromonas gingivalis  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T28651  
R:Han, N.; Whitlock, J.; Proguiske-Fox, A.  
Infect. Immun. 64, 4000-4007, 1996  
A:Title: The hemagglutinin gene A (haga) of Porphyromonas gingivalis 381 contains fou  
A:Reference number: Z20494; MUID:96406248  
A:Accession: T28651  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2628 <HAN>  
A:Cross-references: EMBL:U41807; NID:g1552410; PID:g1469916; PIDN:AAB17128.1  
C:Genetics:  
A:Gene: haga

Query Match 71.0%; Score 49; DB 2; Length 2628;  
Best Local Similarity 66.7%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 949 NGTPNPMPNPMP 960

RESULT 15  
T12203  
transcription factor Vp1 - common ice plant  
C:Species: Mesembryanthemum crystallinum (common ice plant)  
C>Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 20-Jun-2000  
C:Accession: T12203  
R:Fukuhara, T.; Bohnert, H.J.  
submitted to the EMBL Data Library, June 1998  
A:Description: Expression of Vp1 and water channel proteins during seed germination.  
A:Reference number: Z17452  
A:Accession: T12203  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-790 <FTK>  
A:Cross-references: EMBL:AB015183  
C:Genetics:  
A:Gene: Vp1  
C:Superfamily: rice transcription factor Vp1  
C:Keywords: DNA binding; transcription regulation

Query Match 69.6%; Score 48; DB 2; Length 790;  
Best Local Similarity 72.7%; Pred. No. 6.1;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ANPNANPNANP 12  
|:| | | | |  
Db 377 ADPNPNPNPNP 387

Search completed: January 29, 2002, 10:26:35  
Job time: 646 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:37 ; Search time 80.65 seconds  
(without alignments)  
5.455 Million cell updates/sec

Title: US-09-763-397A-4  
Perfect score: 69  
Sequence: 1 NANPNANPNANP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues  
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	69	100.0	315	1	CSP_PLAFL	P05691 plasmodium
2	69	100.0	388	1	CSP_PLARE	P26694 plasmodium
3	69	100.0	397	1	CSP_PLAFO	P19597 plasmodium
4	69	100.0	412	1	CSP_PLAFA	P02893 plasmodium
5	69	100.0	424	1	CSP_PLAFT	P13814 plasmodium
6	69	100.0	442	1	CSP_PLAFW	P08307 plasmodium
7	49	71.0	2628	1	HAGA_PORGI	Q51845 porphyromon
8	43	62.3	441	1	GUN2_THEFU	P26222 thermomonos
9	41	59.4	162	1	CRA_PLAFA	P04923 plasmodium
10	41	59.4	162	1	EXP1_PLAFA	P04926 plasmodium
11	41	59.4	463	1	PLSB_CARTI	Q42713 carthamus t
12	41	59.4	1314	1	ADR6_YEAST	P09547 saccharomyc
13	40	58.0	64	1	CYPM_BOVIN	P30404 bos taurus
14	40	58.0	169	1	SNFB_YEAST	P38956 saccharomyc
15	40	58.0	365	1	PROB_BACSU	P39820 bacillus su
16	40	58.0	438	1	DCDA_METJA	Q58497 methanococc
17	40	58.0	766	1	EVA_DROME	Q05201 drosophila
18	40	58.0	3828	1	TRX_DROVI	Q24742 drosophila
19	39	56.5	430	1	SGK_RAT	Q06226 rattus norv
20	39	56.5	431	1	SGK_HUMAN	Q00141 homo sapien
21	39	56.5	431	1	SGK_MOUSE	Q9wvc6 mus musculu
22	39	56.5	431	1	SGK_RABIT	Q9xtl8 oryctolagus
23	39	56.5	484	1	P60_LISMO	P21171 listeria mo
24	39	56.5	490	1	MOT3_YEAST	P54785 saccharomyc
25	39	56.5	672	1	NGR1_YEAST	P32831 saccharomyc
26	39	56.5	919	1	VGLB_HSVB4	P17472 equine herp
27	39	56.5	1531	1	NFT5_HUMAN	O94916 homo sapien
28	39	56.5	2410	1	MOK1_SCHPO	Q9usk8 schizosacch
29	38.5	55.8	848	1	AMYB_ARATH	P25853 arabidopsis
30	38.5	55.8	844	1	GCRL_YEAST	P07261 saccharomyc
31	38	55.1	339	1	CSP_PLABE	P06915 plasmodium
32	38	55.1	347	1	CSP_PLABA	P23093 plasmodium
33	38	55.1	402	1	PROB_SCHPO	O13810 schizosacch

RESULT 1

ID	CSP_PLAFL	STANDARD	PRT	315 AA
AC	P05691;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	01-NOV-1988 (Rel. 09, Last sequence update)			
DT	01-FEB-1994 (Rel. 28, Last annotation update)			
DE	CIRCUMSPOROZOITE PROTEIN (CS) (FRAGMENT).			
OS	Plasmodium falciparum (isolate le5).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5840;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-87308186; PubMed-2442154;			
RA	la Cruz V.F., Lal A.A., McCutchan T.F.;			
RT	"Sequence variation in putative functional domains of the			
RT	circumsporozoite protein of Plasmodium falciparum. Implications for			
RT	vaccine development."			
RL	J. Biol. Chem. 262:11935-11939(1987).			
CC	-!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT			
CC	SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE			
CC	MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE			
CC	VERTEBRATE HOST).			
CC	-!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR			
CC	ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES			
CC	WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M17802; AAA29538.1; -			
KW	Malaria; Sporozoite; Repeat.			
FT	NON_TER 1			
FT	DOMAIN 107 270 40 X 4 AA TANDEM REPEATS OF P-N-A-N.			
FT	NON_TER 315 315			
SQ	SEQUENCE 315 AA; 33649 MW; A334DB11FA7FD777 CRC64;			

Query Match 100.0%; Score 69; DB 1; Length 315;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
|||||  
DB 132 NANPNANPNANP 143

RESULT 2

ID	CSP_PLARE	STANDARD	PRT	388 AA
AC	P26694;			

5

```
CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CC NCBI_TaxID=5833;
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=84250215; PubMed=6204383;
RA Dane J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hochmeyer W.F., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984)
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02194; AAA29524.1; -.
DR PIR; A03388; OZZQAF.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 412
FT DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
SQ SEQUENCE 412 AA; 44420 MW; 1EEED3DE90965F8 CRC64;

Query Match 100.0%; Score 69; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12
Db 148 NANPNANPNANP 159
|||||
CSP_PLAFT STANDARD; PRT; 424 AA.

RESULT 5
CSP_PLAFT
ID CSP_PLAFT STANDARD; PRT; 424 AA.
AC P13814;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium falciparum (isolate t4 / Thailand).
CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CC NCBI_TaxID=5846;
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzeiwig R.S., Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from
RT Thailand.";
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -----
```

```
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19752; AAA29555.1; -.
DR PIR; A54533; A54533.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 424
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;

Query Match 100.0%; Score 69; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12
Db 148 NANPNANPNANP 159
|||||
CSP_PLAFT STANDARD; PRT; 442 AA.

RESULT 6
CSP_PLAFT
ID CSP_PLAFT STANDARD; PRT; 442 AA.
AC P08307;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium falciparum (isolate Wellcome).
CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CC NCBI_TaxID=5848;
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=87115616; PubMed=3543671;
RA Lockyer M.J., Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium
RT falciparum.";
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M15505; AAA29554.1; -.
DR PIR; A54529; A54529.
```

DR	PROSITE; PS00656; GLYCOSYL_HYDROL_F6_2; 1.
KW	Cellulose degradation; Hydrolase; Glycosidase; Signal; 3D-structure.
FT	SIGNAL 1 31
FT	CHAIN 32 441
FT	ENDOGLUCANASE E-2.



```
FT DOMAIN 320 CATALYTIC.
FT DOMAIN 321 LINKER.
FT DOMAIN 340 CELLULOSE-BINDING.
FT ACT_SITE 110
FT ACT_SITE 148
FT ACT_SITE 296
FT DISULFID 111
FT DISULFID 263
FT DISULFID 346
FT STRAND 36
FT TURN 40
FT HELIX 43
FT TURN 52
FT TURN 53
FT TURN 55
FT HELIX 57
FT TURN 64
FT STRAND 69
FT STRAND 71
FT HELIX 78
FT TURN 96
FT STRAND 98
FT STRAND 100
FT TURN 109
FT HELIX 122
FT TURN 135
FT STRAND 142
FT TURN 147
FT HELIX 149
FT TURN 154
FT HELIX 158
FT TURN 180
FT STRAND 183
FT HELIX 196
FT TURN 206
FT HELIX 208
FT STRAND 214
FT TURN 219
FT TURN 220
FT TURN 225
FT TURN 240
FT TURN 242
FT STRAND 245
FT TURN 254
FT TURN 259
FT TURN 266
FT STRAND 275
FT TURN 281
FT STRAND 283
FT TURN 292
FT STRAND 294
FT TURN 302
FT STRAND 305
FT HELIX 307
FT TURN 316
SQ SEQUENCE 441 AA; 87218E4537092AE5 CRC64;
```

Query Match 62.3%; Score 43; DB 1; Length 441;  
Best Local Similarity 70.0%; Pred. No. 11;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy 3 NPANPNANP 12
    ||| ||| |
Db 321 NPANPNPTP 330
```

RESULT 9

```
ID CRA_PLAFA STANDARD; PRT; 162 AA.
AC P04923;
```

```
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CIRCUMSPOROZOITE-PROTEIN RELATED ANTIGEN PRECURSOR (CRA).
```

```
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270476; PubMed=2410913;
RA Coppel R.L., Favaloro J.M., Crewther P.E., Burkot T.R., Bianco A.E.,
RA Stahl H.-D., Kemp D.J., Anders R.F., Brown G.V.;
RT "A blood stage antigen of Plasmodium falciparum shares determinants
RL with the sporozoite coat protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:5121-5125(1985).
CC -!- MISCELLANEOUS: ALTHOUGH THERE ARE NO AUTHENTIC REPEATS IN THIS
CC ANTIGEN, THERE ARE A NUMBER OF INTERNAL HOMOLOGIES (N-A-N-P) AND
CC (N-A-D-P), THE FIRST OF THESE TETRAMERS IS THE DOMINANT REPEAT
CC FOUND IN THE CSP OF P. FALCIPARUM & REACTS WITH ANTIBODIES AGAINST
CC CRA. IT IS POSSIBLE THAT IMMUNE RESPONSES TO CRA MAY ACT AGAINST
CC THE CSP ALSO. THE CRA IS FOUND IN MANY PARASITIC STRAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11145; AAA29523.1; .
DR PIR; A25780; A25780.
KW Signal; Malaria.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 162 CIRCUMSPOROZOITE-PROTEIN RELATED ANTIGEN.
SQ SEQUENCE 162 AA; 17349 MW; FB60C8250BC0B589 CRC64;
```

Query Match 59.4%; Score 41; DB 1; Length 162;

Best Local Similarity 50.0%; Pred. No. 8.1;

Matches 9; Conservative 1; Mismatches 2; Indels 6; Gaps 1;

```
Qy 1 NANP-----NANPNANP 12
    |||| | ||| |
Db 120 NANPDADSESGEPNADP 137
```

RESULT 10

```
EXPL_PLAFA STANDARD; PRT; 162 AA.
ID AC P04926; P06718;
```

```
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MALARIA PROTEIN EXP-1 PRECURSOR (EXPORTED ANTIGEN AG 5.1).
```

GN EXP-1.

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=5833;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87218504; PubMed=2438130;

RA Simmons D., Woollett G., Bergin-Cartwright M., Kay D., Scaife J.;

RT "A malaria protein exported into a new compartment within the host erythrocyte.";

RL EMBO J. 6:485-491(1987).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=85215483; PubMed=2582354;

RA Hope I.A., McKay M., Hyde J.E., Goman M., Scaife J.;

RT "The gene for an exported antigen of the malaria parasite Plasmodium falciparum cloned and expressed in Escherichia coli.";

RL Nucleic Acids Res. 13:369-379(1985).

CC -!- MISCELLANEOUS: THIS PROTEIN IS ADDRESSED TO A NEW COMPARTMENT

CC WITHIN THE CYTOPLASM OF THE INFECTED RED CELL. IT ALSO SURROUNDS

CC THE PARASITE, PROBABLY IN THE PARASITOPHOROUS VACUOLE MEMBRANE.

CC -!- MISCELLANEOUS: THIS ANTIGEN AND THE CIRCUMSPOROZOITE PROTEIN



```
KW Zinc-finger. 5 65 ASN/THR-RICH.
FT DOMAIN 337 385 GLN-RICH.
FT ZN_FING 1241 1258 C4-TYPE.
SQ SEQUENCE 1314 AA; 147938 MW; F442D5A82013CDBD CRC64;

Query Match 59.4%; Score 41; DB 1; Length 1314;
Best Local Similarity 63.6%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NANPNANPNAN 11
Db 223 NTNSNSTPNAN 233

RESULT 13
CYPM_BOVIN STANDARD; PRT; 64 AA.
AC P30404;
DT 01-APR-1993 (Rel. 25, Last Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PPTIDYL-PROLYL CIS-TRANS ISOMERASE, MITOCHONDRIAL (EC 5.2.1.8)
DE (PPIASE) (ROTAMASE) (CYCLOPHILIN F) (FRAGMENTS).
GN PPIF OR CYP3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
ON NCBI_TaxID=9913;
RX MEDLINE=93176190; PubMed=7679902;
RC TISSUE=Heart;
RA Inoue T., Yoshida Y., Isaka Y., Tagawa K.;
RT "Isolation of mitochondrial cyclophilin from bovine heart.";
RL Biochem. Biophys. Res. Commun. 190:857-863(1993).
CC -!- FUNCTION: PPIASES ACCELERATE THE FOLDING OF PROTEINS.
CC -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC PEPTIDE BONDS IN OLIGOPEPTIDES.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -!- SIMILARITY: BELONGS TO THE CYCLOPHILIN-TYPE PPIASE FAMILY.
DR PIR; PC1237; PC1237.
DR InterPro: IPR002130; CSA_PPIase.
DR Pfam: PF00160; PRO_isomerase; 1.
DR PROSITE; PS00170; CSA_PPIASE_1; PARTIAL.
DR PROSITE; PS0072; CSA_PPIASE_2; PARTIAL.
KW Cyclosporin; Isomerase; Rotamase; Multigene family; Mitochondrion.
FT NON_CONS 20 21
FT NON_CONS 30 31
FT NON_CONS 50 51
FT NON_TER 64 64
SQ SEQUENCE 64 AA; 6472 MW; 80926AF5625B9EE5 CRC64;

Query Match 58.0%; Score 40; DB 1; Length 64;
Best Local Similarity 63.6%; Pred. No. 4.5;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NANPNANPNAN 11
Db 28 NGNANAGPNTN 38

RESULT 14
SNFB_YEAST STANDARD; PRT; 169 AA.
AC P38956;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TRANSCRIPTION REGULATORY PROTEIN SNF1 (SWI/SNF COMPLEX COMPONENT
```

```
DE SNF1).
KW SNF1 OR YDR073W OR D4411.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RA Coster F., Jonniaux J.-L., Goffeau A.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=95349590; PubMed=7623818;
RA Treich I., Cairns B.R., de Los Santos T., Brewster E., Carlson M.;
RT "SNF1, a new component of the yeast SNF-SWI complex that interacts
RT with a conserved region of SNF2.";
RL Mol. Cell. Biol. 15:4240-4248(1995).
CC -!- FUNCTION: INVOLVED IN TRANSCRIPTIONAL ACTIVATION. THE SWI/SNF
CC COMPLEX IS REQUIRED FOR THE INDUCED EXPRESSION OF A LARGE NUMBER
CC OF GENES. THIS COMPLEX ALTERS CHROMATIN STRUCTURE TO FACILITATE
CC BINDING OF GENE-SPECIFIC DEDICATED TRANSCRIPTION FACTORS.
CC -!- SUBUNIT: COMPONENT OF THE SWI/SNF GLOBAL TRANSCRIPTION ACTIVATOR
CC COMPLEX.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46796; CAA86795.1; -
DR EMBL; Z74369; CAA98891.1; -
DR EMBL; X82086; CAA57601.1; -
DR EMBL; X82086; CAA57600.1; -
DR PIR; S48760; S48760.
DR SGD; S0002480; SNF1.
KW Transcription regulation; Activator; Nuclear protein; Repeat.
FT DOMAIN 28 51 6 X 4 AA TANDEM REPEATS OF N-[AT]-[NT])-A.
FT REPEAT 28 31 1-1.
FT REPEAT 32 35 1-2.
FT REPEAT 36 39 1-3.
FT REPEAT 40 43 1-4.
FT REPEAT 44 47 1-5.
FT REPEAT 48 51 1-6.
FT DOMAIN 76 165 2 X 5 AA REPEATS OF L-L-A-R-V.
FT REPEAT 76 80 2-1.
FT REPEAT 160 165 2-2.
FT DOMAIN 93 102 POLY-ASN.
SQ SEQUENCE 169 AA; 18667 MW; FC626E373C32C2C7 CRC64;

Query Match 58.0%; Score 40; DB 1; Length 169;
Best Local Similarity 72.7%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NANPNANPNAN 11
Db 30 NANANANATAN 40

RESULT 15
PROB_BACSU STANDARD; PRT; 365 AA.
ID PROB_BACSU
AC P39820; O34562;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GLUTAMATE 5-KINASE (EC 2.7.2.11) (GAMMA-GLUTAMYL KINASE) (GK).
GN PROB.
```

OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.  
RX STRAIN=168;  
RA MEDLINE=94364946; PubMed=8083159;  
RA Ogura M., Kawata-Mukai M., Itaya M., Takio K., Tanaka T.;  
RT "Multiple copies of the proB gene enhance degS-dependent  
RT extracellular protease production in Bacillus subtilis."  
RL J. Bacteriol. 176:5673-5680(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=168;  
RA Devine K.M.;  
RL "Sequence of the Bacillus subtilis genome between xlyA and ykoR";  
CC Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE = ADP + L-GLUTAMATE  
CC 5-PHOSPHATE (PRODUCT RAPIDLY CYCLIZES TO 5-OXOPROLINE AND  
CC ORTHOPHOSPHATE).  
CC -1- PATHWAY: FIRST STEP IN PROLINE BIOSYNTHESIS PATHWAY.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE GLUTAMATE 5-KINASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D26044; BAA05044.1; -  
DR EMBL; AJ002571; CAA05591.1; -  
DR EMBL; Z99110; CAB13169.1; -  
DR Subtilist; BG10963; proB.  
DR InterPro; IPR001048; AakKinase.  
DR InterPro; IPR001057; Glut\_5\_kinase.  
DR Pfam; PF00696; aakKinase; 1.  
DR Pfam; PF01472; PUA; 1.  
DR PRINTS; PR00474; GLU5KINASE.  
DR SMART; SM00359; PUA; 1.  
DR PROSITE; PS00902; GLUTAMATE\_5\_KINASE; 1.  
KW Transferase; Kinase; Proline biosynthesis; Complete proteome.  
FT CONFLICT 25 32 AKIRHVQ -> ONQRACS (IN REF. 1).  
FT CONFLICT 51 51 P -> A (IN REF. 1).  
FT CONFLICT 54 54 R -> A (IN REF. 1).  
FT CONFLICT 181 185 NPEAK -> ILSE (IN REF. 1).  
SQ SEQUENCE 365 AA; 39457 MW; 042871624A0A7291 CRC64;

Query Match 58.0%; Score 40; DB 1; Length 365;  
Best Local Similarity 70.0%; Pred. No. 25;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NANPNANPNA 10  
:|||||  
DB 175 DANPNENPEA 184

Search completed: January 29, 2002, 11:13:38  
Job time: 814 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:09 ; Search time 285.36 Seconds  
(without alignments)  
6.151 Million cell updates/sec

Title: US-09-763-397A-4  
Perfect score: 69  
Sequence: 1 NANPNANPNANP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 segs, 146272329 residues  
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_17.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	69	100.0	383	5	Q9GPN1	Q9gpn1 plasmodium
2	69	100.0	393	5	Q99255	Q99255 plasmodium
3	69	100.0	408	5	Q25729	Q25729 plasmodium
4	69	100.0	416	5	Q25829	Q25829 plasmodium
5	69	100.0	420	5	Q25831	Q25831 plasmodium
6	69	100.0	420	5	Q25838	Q25838 plasmodium
7	69	100.0	424	5	Q27425	Q27425 plasmodium
8	69	100.0	424	5	Q99256	Q99256 plasmodium
9	69	100.0	432	5	Q27246	Q27246 plasmodium
10	69	100.0	432	5	Q25827	Q25827 plasmodium
11	69	100.0	436	5	Q27325	Q27325 plasmodium
12	69	100.0	436	5	Q25828	Q25828 plasmodium
13	69	100.0	442	5	Q25830	Q25830 plasmodium
14	69	100.0	452	5	Q25834	Q25834 plasmodium
15	57	82.6	297	5	Q9V650	Q9v650 drosophila
16	57	82.6	299	5	O61464	O61464 drosophila
17	56	81.2	378	5	Q9V8Q7	Q9v8q7 drosophila
18	55	79.7	502	5	Q9W4T4	Q9w4t4 drosophila
19	55	79.7	530	5	O76918	O76918 drosophila

20	54	78.3	126	12	Q9DH26	Q9dh26 meleagrid h
21	54	78.3	146	10	Q9FY84	Q9fy84 arabidopsis
22	54	78.3	314	11	O54817	O54817 mus musculus
23	54	78.3	382	11	O88933	O88933 mus musculus
24	54	78.3	387	10	Q9M9V1	Q9m9v1 arabidopsis
25	54	78.3	387	11	Q9Z2U3	Q9z2u3 mus musculus
26	54	78.3	1687	2	Q9R9B7	Q9r9b7 porphyromon
27	54	78.3	1706	2	Q51838	Q51838 porphyromon
28	54	78.3	1706	2	Q51839	Q51839 porphyromon
29	53	76.8	214	5	Q9V9Y2	Q9v9y2 drosophila
30	52	75.4	1408	5	Q9W0C9	Q9w0c9 drosophila
31	51	73.9	1231	5	Q97140	Q97140 dictyostell
32	50	72.5	1171	5	O61732	O61732 drosophila
33	50	72.5	1178	5	Q9W3E2	Q9w3e2 drosophila
34	49	71.0	178	10	Q9LRJ8	Q9lrj8 arabidopsis
35	49	71.0	284	10	Q9LJ57	Q9lj57 arabidopsis
36	49	71.0	296	5	Q9VTQ0	Q9vtq0 drosophila
37	49	71.0	417	10	Q9FUK8	Q9fuk8 pharbitis n
38	49	71.0	1097	2	P72196	P72196 porphyromon
39	49	71.0	1223	2	Q9ZNB5	Q9znb5 porphyromon
40	49	71.0	1358	2	P96967	P96967 porphyromon
41	49	71.0	1704	2	Q51816	Q51816 porphyromon
42	49	71.0	1723	2	P72197	P72197 porphyromon
43	49	71.0	1723	2	P72194	P72194 porphyromon
44	49	71.0	1732	2	Q51817	Q51817 porphyromon
45	49	71.0	1732	2	O52050	O52050 porphyromon

ALIGNMENTS

RESULT 1					
ID Q9GPN1	PRELIMINARY;	PRT;	383 AA.		
AC Q9GPN1;					
DT 01-MAR-2001 (Tremblrel. 16, Created)					
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)					
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)					
DE CIRCUMSPOROZITE PROTEIN (FRAGMENT).					
OS Plasmodium falciparum.					
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.					
OX NCBI_TaxID=5833;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=FCC-1/HN;					
RA Zheng C., Xie P., Chen Y.;					
RT "Molecular cloning and expression of circumsporoite protein gene					
RL from Plasmodium falciparum FCC-1/HN strain in mycobacterium.";					
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.					
DR EMBL; AF315469; AAG37074.1; .					
DR InterPro; IPR003067; Circmsprzoite.					
DR InterPro; IPR000884; TSPL.					
DR Pfam; PF00090; tsp_1; 1.					
DR PRINTS; PR01303; CRCMSPRZOITE.					
FT NON_TER 1					
FT NON_TER 383					
SQ SEQUENCE 383 AA; 40893 MW; 503C5DFDF61A9E27 CRC64;					

Query Match 100.0%; Score 69; DB 5; Length 383;  
Best Local Similarity 100.0%; Pred. No. 0.0018;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
|||||  
Db 129 NANPNANPNANP 140

RESULT 2					
Q99255	PRELIMINARY;	PRT;	393 AA.		
ID Q99255					
AC Q99255;					
DT 01-NOV-1996 (Tremblrel. 01, Created)					



RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;  
 RT "Structure of the gene encoding the immunodominant surface antigen on  
 RL the sporozoite of the human malaria parasite Plasmodium falciparum."  
 RN Science 225:593-599(1984).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K1;  
 RA la Cruz V.F.;  
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K1;  
 RA Jongwutiwes S., Tanabe K., Kanbara H.;  
 RL Mol. Biochem. Parasitol. 0:0-0(0).  
 DR EMBL; M831174; AAA29552.1; -;  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR003067; Crmsprzoite.  
 DR Pfam; PF00090; tsp\_1; 1.  
 DR PRINTS; PR01303; CRCMSPRZOITE.  
 DR PROSITE; PS00092; TSP1; 1.  
 DR SMART; SM00209; TSP1; 1.  
 SQ SEQUENCE 420 AA; 45318 MW; F7F70F1C4939DEA7 CRC64;

Query Match 100.0%; Score 69; DB 5; Length 420;  
 Best Local Similarity 100.0%; Pred. No. 0.0019;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 |||||  
 DB 148 NANPNANPNANP 159

## RESULT 6

Q25838  
 ID Q25838 PRELIMINARY; PRT; 420 AA.  
 AC Q25838;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CIRCUMSPOROZOITE PROTEIN.  
 GN CSP.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=835B;  
 RX MEDLINE=84250215; PubMed=6204383;  
 RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
 RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
 RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;  
 RT "Structure of the gene encoding the immunodominant surface antigen on  
 RL the sporozoite of the human malaria parasite Plasmodium falciparum."  
 RN Science 225:593-599(1984).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=835B;  
 RA la Cruz V.F.;  
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=835B;  
 RA Jongwutiwes S., Tanabe K., Kanbara H.;  
 RL Mol. Biochem. Parasitol. 0:0-0(0).  
 DR EMBL; M831161; AAA29574.1; -;  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR003067; Crmsprzoite.  
 DR Pfam; PF00090; tsp\_1; 1.  
 DR PRINTS; PR01303; CRCMSPRZOITE.  
 DR SMART; SM00209; TSP1; 1.  
 SQ SEQUENCE 420 AA; 45155 MW; 3A85B92432C2893C CRC64;

Query Match 100.0%; Score 69; DB 5; Length 420;  
 Best Local Similarity 100.0%; Pred. No. 0.0019;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 |||||  
 DB 148 NANPNANPNANP 159

## RESULT 7

Q27425  
 ID Q27425 PRELIMINARY; PRT; 424 AA.  
 AC Q27425;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CIRCUMSPOROZOITE PROTEIN.  
 GN CSP.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=844;  
 RX MEDLINE=84250215; PubMed=6204383;  
 RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
 RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
 RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;  
 RT "Structure of the gene encoding the immunodominant surface antigen on  
 RL the sporozoite of the human malaria parasite Plasmodium falciparum."  
 RN Science 225:593-599(1984).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=844;  
 RA la Cruz V.F.;  
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=844;  
 RA Jongwutiwes S., Tanabe K., Kanbara H.;  
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M831169; AAA29547.1; -;  
 DR EMBL; M831149; AAA29562.1; -;  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR003067; Crmsprzoite.  
 DR Pfam; PF00090; tsp\_1; 1.  
 DR PRINTS; PR01303; CRCMSPRZOITE.  
 DR SMART; SM00209; TSP1; 1.  
 SQ SEQUENCE 424 AA; 45592 MW; F20CEB60636DB98E CRC64;

Query Match 100.0%; Score 69; DB 5; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 0.0019;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
 |||||  
 DB 148 NANPNANPNANP 159

## RESULT 8

Q99256  
 ID Q99256 PRELIMINARY; PRT; 424 AA.  
 AC Q99256;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CIRCUMSPOROZOITE PROTEIN (CS) PRECURSOR, VARIANT 2 (FRAGMENT).  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]

RC  
PC  
RA  
RL  
DR  
DR  
DR  
DR  
DR  
DR

RESULT	11		
Q27325			
ID	Q27325	PRELIMINARY;	PRT; 436 AA.
AC	Q27325;		
DT	01-NOV-1996	(TREMBlrel. 01, Created)	



DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN.  
GN CSP.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=837;  
RX MEDLINE=84250215; PubMed=6204383;  
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
RA Sanders G.S., Reddy P.E., Digs C.L., Miller L.H.;  
RT "Structure of the gene encoding the immunodominant surface antigen on  
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";  
RL Science 225:593-599(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=837;  
RC Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=837;  
RC Jongwutiwes S., Tanabe K., Kanbara H.;  
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
DR EMBL; M83164; AAA29542.1; -;  
DR EMBL; M83150; AAA29563.1; -;  
DR EMBL; M83163; AAA29576.1; -;  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR003067; Crmsprzoite.  
DR Pfam; PF00090; tsp\_1; 1.  
DR PRINTS; PR01303; CRCMSPRZOITE.  
DR SMART; SM00209; TSP1; 1.  
SQ SEQUENCE 436 AA; 4668 MW; 5842FF3348B68655 CRC64;

Query Match 100.0%; Score 69; DB 5; Length 436;  
Best Local Similarity 100.0%; Pred. No. 0.002;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
Db 140 NANPNANPNANP 151  
|||||  
|||||

RESULT 12  
Q25828  
ID Q25828 PRELIMINARY; PRT; 436 AA.  
AC Q25828;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN.  
GN CSP.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=842;  
RC Jongwutiwes S., Tanabe K., Kanbara H.;  
RL Mol. Biochem. Parasitol. 0:0-0(0).  
DR EMBL; M83173; AAA29551.1; -;  
DR InterPro; IPR000884; TSP1.  
DR Pfam; PF00090; tsp\_1; 1.  
DR PRINTS; PR01303; CRCMSPRZOITE.  
DR SMART; SM00209; TSP1; 1.  
SQ SEQUENCE 442 AA; 47414 MW; BFAF9D939D7862FF CRC64;

Query Match 100.0%; Score 69; DB 5; Length 442;  
Best Local Similarity 100.0%; Pred. No. 0.002;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
Db 150 NANPNANPNANP 161  
|||||  
|||||

RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=842;  
RC Jongwutiwes S., Tanabe K., Kanbara H.;  
RL Mol. Biochem. Parasitol. 0:0-0(0).  
DR EMBL; M83167; AAA29545.1; -;  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR003067; Crmsprzoite.  
DR Pfam; PF00090; tsp\_1; 1.  
DR PRINTS; PR01303; CRCMSPRZOITE.  
DR PROSITE; PS00092; TSP1; 1.  
DR SMART; SM00209; TSP1; 1.  
SQ SEQUENCE 436 AA; 46875 MW; F102683C5C1DC85A CRC64;

Query Match 100.0%; Score 69; DB 5; Length 436;  
Best Local Similarity 100.0%; Pred. No. 0.002;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
Db 148 NANPNANPNANP 159  
|||||  
|||||

RESULT 13  
Q25830  
ID Q25830 PRELIMINARY; PRT; 442 AA.  
AC Q25830;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN.  
GN CSP.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TQ/94;  
RC MEDLINE=84250215; PubMed=6204383;  
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
RA Sanders G.S., Reddy P.E., Digs C.L., Miller L.H.;  
RT "Structure of the gene encoding the immunodominant surface antigen on  
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";  
RL Science 225:593-599(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TQ/94;  
RC la Cruz V.F.;  
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TQ/94;  
RC Jongwutiwes S., Tanabe K., Kanbara H.;  
RL Mol. Biochem. Parasitol. 0:0-0(0).  
DR EMBL; M83173; AAA29551.1; -;  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR003067; Crmsprzoite.  
DR Pfam; PF00090; tsp\_1; 1.  
DR PRINTS; PR01303; CRCMSPRZOITE.  
DR SMART; SM00209; TSP1; 1.  
SQ SEQUENCE 442 AA; 47414 MW; BFAF9D939D7862FF CRC64;

Query Match 100.0%; Score 69; DB 5; Length 442;  
Best Local Similarity 100.0%; Pred. No. 0.002;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12  
Db 150 NANPNANPNANP 161  
|||||  
|||||

RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA	Burtis J.C., Busam D.A., Butler H., Cadieu E., Centor A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA	Spter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT	"The genome sequence of Drosophila melanogaster."
RL	Science 287:2185-2195(2000).
DR	EMBL; AE003823; AAF58589.1;
DR	FlyBase; FBgn0024732; Rep1.
DR	InterPro; IPR003508; CAD.
DR	Pfam; PF02017; CIDE-N; 1.
DR	SMART; SM00266; CAD; 1.
SQ	SEQUENCE 297 AA; 31985 MW; 231804A8E47F5606 CRC64;
QY	1 NANPNANPNAN 11
DB	284 NANPNANPNAN 294
Query Match	82.6%; Score 57; DB 5; Length 297;
Best Local Similarity	90.9%; Pred. No. 0.099;
Matches	10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 NANPNANPNAN 11
DB	284 NANPNANPNAN 294
Search completed:	January 29, 2002, 11:12:09
Job time:	765 sec

RA	Q25834; PRELIMINARY; PRT; 452 AA.
AC	Q25834; PRELIMINARY; PRT; 452 AA.
DT	01-NOV-1996 (TREMBLrel. 01, Created)
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	CIRCUMSPOROZOITE PROTEIN.
GN	CSP.
OS	Plasmodium falciparum.
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=5833;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=827;
RC	MEDLINE=84250215; PubMed=6204383;
RA	Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA	Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA	Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT	"Structure of the gene encoding the immunodominant surface antigen on
RT	the sporozoite of the human malaria parasite Plasmodium falciparum."
RL	Science 225:593-599(1984).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=827;
RA	la Cruz V.F.;
RL	Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=827;
RC	Jongwutives S., Tanabe K., Kanbara H.;
RL	Mol. Biochem. Parasitol. 0:0-0(0).
DR	EMBL; M83156; AAA29569.1;
DR	InterPro; IPR000884; TSPI.
DR	InterPro; IPR003067; Circumsporozoite.
DR	Pfam; PF00090; tsp_1; 1.
DR	PRINTS; PR01303; CRCMSPRZOITE.
DR	PROSITE; PS00092; TSPI; 1.
DR	SMART; SM00209; TSPI; 1.
SQ	SEQUENCE 452 AA; 48431 MW; 6E739D6C53223805 CRC64;
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NANPNANPNANP 12
DB	140 NANPNANPNANP 151
Query Match	100.0%; Score 69; DB 5; Length 452;
Best Local Similarity	100.0%;



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:21:41 ; Search time 310.82 Seconds  
(without alignments)  
3.098 Million cell updates/sec

Title: US-09-763-397A-5  
Perfect score: 75  
Sequence: 1 KHKLLKQPGDGNP 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	13	21	Plasmodium falciparum
2	75	100.0	19	9	Region of P.falcip
3	75	100.0	180	11	Circumsporozoite a
4	75	100.0	184	11	Circumsporozoite a
5	75	100.0	250	7	Synthetic peptide
6	75	100.0	309	12	NS1_81-Rlfdelta9
7	75	100.0	319	11	NS18RlfAuth plasm
8	75	100.0	319	12	NS1_81-RlfAuth. p
9	75	100.0	327	12	NS1_81-RlfAuth + (
10	75	100.0	335	12	NS1_81(NANP)4RLfau
11	75	100.0	335	12	NS1_81(NVDP)4RLfau

12	75	100.0	350	21	AA70278	Recombinant vaccin
13	75	100.0	411	9	AA83144	Sequence encoded b
14	75	100.0	412	7	AA60416	CS protein of mala
15	75	100.0	412	9	AA80835	Sequence encoded b
16	69	92.0	20	16	AA71653	CS Region I mimeti
17	56	74.7	14	7	AA61001	Sequence corresp.
18	47	62.7	87	21	AA83167	PAGE-4 polypeptide
19	47	62.7	102	21	AA83158	PAGE-4 polypeptide
20	47	62.7	102	21	AA52998	Human secreted pro
21	46	61.3	101	9	AA81691	Sequence encoded b
22	46	61.3	101	9	AA80813	Sequence of R prot
23	46	61.3	101	13	AA22369	Sivmac239 vpr gene
24	46	61.3	101	20	AA99813	HIV-2/SIV Vpr prot
25	46	61.3	101	20	AA89317	Sivmac239 genome v
26	45	60.0	8	12	AA812308	IDI plasmodium sur
27	44	58.7	55	21	AA802761	Human secreted pro
28	43	57.3	135	15	AA851001	M24-M5-M6-M19-M3-M
29	43	57.3	135	21	AA803124	S. pyogenes octava
30	41.5	55.3	20	14	AA837974	P. vivax circumspo
31	41.5	55.3	160	11	AA807288	Circumsporozoite a
32	41.5	55.3	378	8	AA70708	Sequence encoding
33	41.5	55.3	378	14	AA830609	Plasmodium vivax c
34	40	53.3	89	20	AA112447	Human 5' EST seque
35	40	53.3	776	11	AA806400	Lambda gt10ch2 enc
36	40	53.3	866	11	AA806379	Lambda gt10ch101 e
37	40	53.3	954	13	AA829770	Porcine PAM-8. Su
38	40	53.3	973	13	AA820026	C-terminal amidati
39	40	53.3	974	11	AA806399	Lambda gt10ch201 e
40	40	53.3	980	12	AA811113	Equine C-terminal
41	40	53.3	1020	13	AA820025	C-terminal amidati
42	39	52.0	17	21	AA826818	Membrane targeted
43	39	52.0	63	21	AA803302	Human secreted pro
44	39	52.0	385	22	AA875195	Drosophila gustato
45	39	52.0	452	19	AA898334	H. pylori GRPO 173

ALIGNMENTS

RESULT 1

AA70281  
ID AA70281 standard; peptide: 13 AA.

AC AA70281;

DT 06-JUN-2000 (first entry)

DE Plasmodium falciparum CSP antigenic epitope, P592.

KW Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;  
KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;  
KW Circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;  
KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;  
KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;  
KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;  
KW Pf927; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody.

OS Plasmodium falciparum.

PN WO200011179-A1.

PD 02-MAR-2000.

PF 19-AUG-1999; 99WO-US18869.

PR 21-AUG-1998; 98US-0097703.

PA (NAIM-) NAT INST IMMUNOLOGY.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Lal AA, Shi YP, Hasnain SE;

DR WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection  
 PT comprises antigenic peptides obtained from different stages of  
 PT plasmodium falciparum life cycle -  
 XX  
 XX Claim 2; Page 16; 52pp; English.  
 PS  
 XX The present sequence is the antigenic epitope p592, derived from  
 CC circumsporozoite protein (CSP) of the sporozoite stage of Plasmodium  
 CC falciparum. It is used in the construction of recombinant protein  
 CC CDC/NIMAVAC-1, which is a multivalent, multistage malarial vaccine.  
 CC The recombinant protein comprises, melittin signal peptide, (His)6 tag,  
 CC T-cell epitope from tetanus toxoid and 21 antigenic epitopes from  
 CC circumsporozoite protein (CSP), sporozoite surface protein-2 (SSP-2),  
 CC liver stage antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1),  
 CC MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding antigen-175  
 CC (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete specific  
 CC antigen, Pf27. These epitopes were obtained at different stages of the  
 CC life cycle of P. falciparum. CDC/NIMAVAC-1 vaccine has antiparasitic  
 CC activity and can be used for treatment and prevention of malarial  
 CC infections. Anti-CDC/NIMAVAC-1 antibodies can be used for detecting  
 CC P. falciparum in biological samples.  
 XX  
 XX Sequence 13 AA;

Query Match 100.0%; Score 75; DB 21; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05; Indels 0; Gaps 0;  
 Matches 13; Conservative 0; Mismatches 0;

QY 1 KKKLKQPGDGNP 13  
 | | | | | | | | | |  
 Db 1 khkkikpgdgndp 13

RESULT 2  
 AAP82541  
 ID AAP82541 standard; peptide; 19 AA.

XX AAP82541;  
 AC  
 XX 12-DEC-1990 (first entry)  
 DT  
 XX Region of P.falciparum circumsporozoite protein with linker attached.  
 DE  
 XX P.falciparum circumsporozoite protein; malaria;  
 KW  
 XX synthetic.

XX DE3723583-A.  
 PN  
 XX 28-JAN-1988.  
 PD  
 XX 16-JUL-1987; 87DE-3723583.  
 PF  
 XX 16-JUL-1986; 86IT-0021144.  
 PR  
 XX (ENTE ) ENIRICERCH SPA.  
 PA

XX Bernardi A, Bonelli F, Pessi A, Verdini AS;  
 PI  
 XX WPI; 1988-030026/05.  
 DR

XX New immunogenic polypeptide for malaria vaccine - contg.  
 PT sequences derived from P. falciparum circumsporozoite protein,  
 PT also useful for antibody assay  
 PT

XX Claim 1; Page 2; 13pp; German.

XX This is the N-terminal of an immunogenic polypeptide useful as a  
 CC malaria vaccine. It comprises a fragment of the P.falciparum  
 CC circumsporozoite protein (CSP) and one copy of a peptide linker,  
 CC which is pref. present in 3-10 copies, which links the preceding

CC sequence to a 2nd (C-terminal) sequence comprising tetrapeptide  
 CC units of CSP. See AAP82542.

SQ Sequence 19 AA;

Query Match 100.0%; Score 75; DB 9; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKLKQPGDGNP 13  
 | | | | | | | | | |  
 Db 3 khkkikpgdgndp 15

RESULT 3  
 AAR07290  
 ID AAR07290 standard; protein; 180 AA.

XX AAR07290;  
 AC  
 XX 28-JAN-1991 (first entry)  
 DT  
 XX Circumsporozoite analogue Falciparum 4.  
 DE

XX CS protein; plasmodium; malaria; vaccine.

XX Synthetic.

XX EP392820-A.

XX 17-OCT-1990.

XX 11-APR-1990; 90EP-0303907.

XX 11-APR-1989; 89US-0336288.

XX (CHIR-) CHIRON CORP.

XX Barr PJ, Bathurst IC, Gibson HL;

XX WPI; 1990-314486/42.

XX N-PSDB; AAQ06153.

XX Recombinant plasmodium circumsporozoite analogues - lacking  
 PT one or more repeat epitope(s) for use as a malaria sub-unit  
 PT vaccine.

XX Claim 10; Fig 10; 22pp; English.

XX The protein was produced by expression of a synthetic gene. The  
 CC analogue comprises AAs 68-123 of the native P. falciparum CS  
 CC protein, followed by four repeat sequences (three "B"s, i.e. NANP)  
 CC and one "A", i.e. NVDP), followed by AAs 289-392 of the native  
 CC protein. Reduction of the immunological dominance of the repeats  
 CC relative to the epitopes in the regions flanking the repeats  
 CC enhances sporozoite neutralising activity.  
 CC See also AAR07287-91.

SQ Sequence 180 AA;

Query Match 100.0%; Score 75; DB 11; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 0.00023;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKLKQPGDGNP 13  
 | | | | | | | | | |  
 Db 46 khkkikpgdgndp 58

RESULT 4  
 AAR07289

ID AAR07289 standard; protein; 184 AA.  
AC AAR07289;  
XX  
XX 28-JAN-1991 (first entry)  
XX  
XX Circumsporozoite analogue Falciparum 3.  
XX  
XX CS protein; plasmodium; malaria; vaccine.  
XX  
XX Synthetic.  
XX  
XX EP392820-A.  
XX  
XX 17-OCT-1990.  
XX  
XX 11-APR-1990; 90EP-0303907.  
XX  
XX 11-APR-1989; 89US-0336288.  
XX  
XX (CHIR-) CHIRON CORP.  
XX  
XX Barr PJ, Bathurst IC, Gibson HL;  
XX  
XX WPI; 1990-314486/42.  
XX  
XX N-PSDB; AAQ06152.  
XX  
XX Recombinant plasmodium circumsporozoite analogues - lacking  
PT one or more repeat epitope(s) for use as a malaria sub-unit  
PT vaccine.  
XX  
XX

PS Claim 10; Fig 8; 22pp; English.  
XX  
XX The protein was produced by expression of a synthetic gene. The  
CC analogue comprises AAs 43-123 of the native P. falciparum CS  
CC protein, followed by four repeat sequences (three "B"s, i.e. NANP)  
CC and one "A", i.e. NVDP), followed by AAs 289-374 of the native  
CC protein. Reduction of the immunological dominance of the repeats  
CC relative to the epitopes in the regions flanking the repeats  
CC enhances sporozoite neutralising activity.  
XX See also AAR07287-91.  
XX

SQ Sequence 184 AA;

Query Match 100.0%; Score 75; DB 11; Length 184;  
Best Local Similarity 100.0%; Pred. No. 0.00023;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KHKKLKQPGDGNP 13  
| | | | | | | | | | | | | | | | | |  
Db 68 khkkikqpgdgnp 80

RESULT 5

AAP60412  
ID AAP60412 standard; protein; 250 AA.  
XX  
XX  
XX AAP60412;  
XX  
XX 13-JUN-1991 (first entry)  
XX  
XX Synthetic peptide antigenic for the malaria parasite.  
XX  
XX Sporozoite; vaccination.  
XX  
XX Plasmodium falciparum.  
XX  
XX Key Location/Qualifiers  
XX Region 18..21  
FT /label= Repeat unit  
FT /note= "Of claim 1"  
FT 22..26

FT /label= Repeat unit  
FT /note= "Of claim 1"

PN EP16410-A.

PD 02-JAN-1986.

XX 24-JUN-1985; 85EP-0107794.

XX 26-JUN-1984; 84US-0624564.

XX (USDC ) US SEC OF COMMERCE.

XX (USGO ) US GOVERNMENT.

XX (USSA ) US SEC OF THE ARMY.

XX McCutchan TF, Dame JB, Williams JL, Schneider I;

XX WPI; 1986-008635/02.

XX New immunologically active pure synthetic peptide(s) - used for  
PT protection against infection by malaria parasite.

XX Claim 7; Page 40-41; 49pp; English.

XX The peptides may be synthesised in pure form and used to generate an  
CC immune response in vaccination against malaria. The featured repeat  
CC units are claimed and must be present in copies of 2-1000, in the  
CC P.falciparum genome, the first is coded for 37 times.  
CC See also AAN60362.

XX SQ Sequence 250 AA;

Query Match 100.0%; Score 75; DB 7; Length 250;

Best Local Similarity 100.0%; Pred. No. 0.00032;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHKKLKQPGDGNP 13

| | | | | | | | | | | | | | | | | |

Db 3 khkkikqpgdgnp 15

RESULT 6

AAR13175  
ID AAR13175 standard; Protein; 309 AA.

XX AAR13175;

XX 29-AUG-1991 (first entry)

XX NSI\_81-RLfdelta9.

XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;  
KW hybrid; Influenza virus; non-structural protein 1; fusion.

XX Plasmodium falciparum.

XX Influenza virus (A/PR/8/34/).

XX Key Location/Qualifiers  
XX Region 1..81

FT /label= N-terminal of NSI

FT /note= "Influenza virus nonstructural protein 1"

FT Peptide 82..87

FT /label= synthetic linker

FT Region 88

FT /label= artifact

FT /note= "see comments"

FT Region 89..193

FT /label= AAs 19-123 of CS protein

FT /note= "Region 1 contg. flanking region less

FT signal sequence"

FT Region 194..309

FT /label= AAs 297-412 of CS protein





```

FT      /note= "Region II flanking region"
XX      EP432965-A.
XX      19-JUN-1991.
XX      06-DEC-1990; 90EP-0313257.
XX      08-DEC-1989; 89US-0447746.
XX      (SMIK ) SMITHKLINE BEECHAM.
XX      (USSA ) US SEC OF THE ARMY.
XX      (BIOM-) BIOMEDICAL RES INST.
XX      Gross MS, Gordon DM, Hollingdale MR;
XX      WPI; 1991-179771/25.
XX      Polypeptide comprising immunogenic determinants from P falciparum
XX      PT - for vaccine against malaria infection in humans.
XX      Example 2; Page 10; 18pp; English.
XX      The polypeptide is prepd. by genetic engineering of genes encoding
XX      the P. falciparum circumsporozoite (CS) protein [Dane et al.,
XX      Science 225 : 593 (1984)], and the influenza virus non-structural
XX      protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
XX      (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX      (NS1_81) is linked via a synthetic sequence to DNA encoding Region
XX      I contg. flanking regionless the 18 AA signal region, which in
XX      turn is fused to DNA encoding Region II-contg. flanking region.
XX      This CS fusion is designated RlfAuth. The Pro residue separating
XX      the Asp (at the C-terminal of the linker) from RlfAuth is an arti-
XX      fact of a filled in BamHI site; the Gly separating Region I and
XX      Region II-contg. CS flanking regions is an artifact of a synthetic
XX      FokI/NotI linker. The peptide can be used in a vaccine for
XX      protection against malaria.
XX      The complete nucleotide and AA sequences are given in EP-304720,
XX      filed May 1, 1990.
XX      See also AAR12306-R12311 and AAR13175-R13179.
XX      SQ Sequence 319 AA;

Query Match 100.0%; Score 75; DB 12; Length 319;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHKKLKQPGDGNP 13
Db 179 khkklkpgdgnp 191

RESULT 9
AAR13177
ID AAR13177 standard; Protein; 327 AA.
XX      AAR13177;
XX      29-AUG-1991 (first entry)
XX      NS1_81-RlfAuth + (NANP)2.
XX      Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
XX      hybrid; influenza virus; non-structural protein 1; fusion.
XX      Plasmodium falciparum.
XX      OS Influenza virus (A/PR/8/34/).
XX      Key Location/Qualifiers
XX      Region 1..81
XX      /label= N-terminal of NS1
XX      /note= "Influenza virus nonstructural protein 1"

```

```

FT      Peptide 82..87
XX      /label= synthetic linker
XX      Region 88
XX      /label= artifact
XX      /note= "see comments"
XX      Region 89..193
XX      /label= AAs 19-123 of CS protein
XX      /note= "Region I contg. flanking region less
XX      signal sequence"
XX      Region 194..201
XX      /label= immunodominant repeat region
XX      /note= "two tetrapeptide repeat units"
XX      Region 202
XX      /label= artifact
XX      /note= "see comments"
XX      Region 203..327
XX      /label= AAs 288-412 of CS protein
XX      /note= "Region II flanking region"
XX      EP432965-A.
XX      19-JUN-1991.
XX      06-DEC-1990; 90EP-0313257.
XX      08-DEC-1989; 89US-0447746.
XX      (SMIK ) SMITHKLINE BEECHAM.
XX      (USSA ) US SEC OF THE ARMY.
XX      (BIOM-) BIOMEDICAL RES INST.
XX      Gross MS, Gordon DM, Hollingdale MR;
XX      WPI; 1991-179771/25.
XX      Polypeptide comprising immunogenic determinants from P falciparum
XX      PT - for vaccine against malaria infection in humans.
XX      Example 3; Page 10; 18pp; English.
XX      The polypeptide is prepd. by genetic engineering of genes encoding
XX      the P. falciparum circumsporozoite (CS) protein [Dane et al.,
XX      Science 225 : 593 (1984)], and the influenza virus non-structural
XX      protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
XX      (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX      (NS1_81) is linked via a synthetic sequence to DNA encoding Region
XX      I contg. flanking region less the 18 AA signal region. This is
XX      linked to a synthetic sequence encoding two repeat units from the
XX      immunodominant region, which in turn is fused to DNA encoding the
XX      Region II-contg. flanking region. The Pro residue separating the
XX      Asp (at the C-terminal of the linker) from the Region I-contg. CS
XX      flanking region is an artifact of a filled-in BamHI site; the Gly
XX      separating the repeat units and the Region II-contg. CS flanking
XX      region is an artifact of a synthetic FokI/NotI linker. The
XX      peptide can be used in a vaccine for protection against malaria.
XX      See also AAR12306-R12311 and AAR13175-R13179.
XX      SQ Sequence 327 AA;

Query Match 100.0%; Score 75; DB 12; Length 327;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHKKLKQPGDGNP 13
Db 179 khkklkpgdgnp 191

RESULT 10
AAR13178
ID AAR13178 standard; Protein; 335 AA.
XX

```

AC	AAR13178;	SQ	Sequence	335 AA;
XX				
DT	29-AUG-1991 (first entry)			
XX			Query Match	100.0%; Score 75; DB 12; Length 335;
DE	NS1_81(NANP)4RLfAuth.		Best Local Similarity	100.0%; Pred No. 0.00043;
XX			Matches	13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
KW	Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;			
XX	hybrid; influenza virus; non-structural protein 1; fusion.			
XX				
OS	Plasmodium falciparum.			
OS	Influenza virus (A/PR/8/34).			
XX				
PH	Key			
FT	Region			
FT	1..81			
FT	/label= N-terminal of NS1			
FT	/note= "Influenza virus nonstructural protein 1"			
FT	82..97			
FT	/label= immunodominant repeat region			
FT	/note= "four tetrapeptide repeat units"			
FT	98..103			
FT	/label= synthetic linker			
FT	104			
FT	/label= artifact			
FT	/note= "see comments"			
FT	105..209			
FT	/label= AAs 19-123 of CS protein			
FT	/note= "Region I contg. flanking region less			
FT	signal sequence"			
FT	210			
FT	/label= artifact			
FT	/note= "see comments"			
FT	211..335			
FT	/label= AAs 288-412 of CS protein			
FT	/note= "Region II flanking region"			
XX				
PN	EP432965-A.			
XX				
PD	19-JUN-1991.			
XX				
PF	06-DEC-1990; 90EP-0313257.			
XX				
PR	08-DEC-1989; 89US-0447746.			
XX				
PA	(SMIK ) SMITHKLINE BEECHAM.			
PA	(USSA ) US SEC OF THE ARMY.			
PA	(BIOM-) BIOMEDICAL RES INST.			
XX				
PI	Gross MS, Gordon DM, Hollingdale MR;			
XX				
DR	WPI; 1991-179771/25.			
XX				
PT	Polypeptide comprising immunogenic determinants from P falciparum			
PT	- for vaccine against malaria infection in humans.			
XX				
PS	Example 4; Page 11; 18pp; English.			
XX				
CC	The polypeptide is prepd. by genetic engineering of genes encoding			
CC	the P. falciparum circumsporozoite (CS) protein [Dane et al.,			
CC	Science 225 : 593 (1984)], and the influenza virus non-structural			
CC	protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845			
CC	(1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1			
CC	(NS1_81) is linked to a synthetic sequence encoding four repeat			
CC	units from the immunodominant region, which in turn is linked via			
CC	a synthetic sequence to DNA encoding Region I contg. flanking			
CC	region less the 18 AA signal region. This is linked to DNA			
CC	encoding Region II-contg. flanking region. The Pro residue sep-			
CC	arating the Asp (at the C-terminal of the linker) from the Region			
CC	I-contg. CS flanking region is an artifact of a filled-in BamHI			
CC	site; the Gly separating the Region I and II-contg. CS flanking			
CC	regions is an artifact of a synthetic FokI/TthIII I linker. The			
CC	peptide can be used in a vaccine for protection against malaria.			
CC	See also AAR12306-R12311 and AAR13175-R13179.			
XX				

AAAR13179

29-AUG-1991 (first entry)

NS1\_81(NVDP)4RLfAuth.

Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;

hybrid; influenza virus; non-structural protein 1; fusion.

Plasmodium falciparum.

Influenza virus (A/PR/8/34).

Key

Region

1..81

/label= N-terminal of NS1

/note= "Influenza virus nonstructural protein 1"

82..97

/label= immunodominant repeat region

/note= "four tetrapeptide repeat units"

98..103

/label= synthetic linker

104

/label= artifact

/note= "see comments"

105..209

/label= AAs 19-123 of CS protein

/note= "Region I contg. flanking region less

signal sequence"

210

/label= artifact

/note= "see comments"

211..335

/label= AAs 288-412 of CS protein

/note= "Region II flanking region"

EP432965-A.

19-JUN-1991.

06-DEC-1990; 90EP-0313257.

08-DEC-1989; 89US-0447746.

(SMIK ) SMITHKLINE BEECHAM.

(USSA ) US SEC OF THE ARMY.

(BIOM-) BIOMEDICAL RES INST.

Gross MS, Gordon DM, Hollingdale MR;

WPI; 1991-179771/25.

Polypeptide comprising immunogenic determinants from P falciparum

- for vaccine against malaria infection in humans.

Example 4; Page 11; 18pp; English.

The polypeptide is prepd. by genetic engineering of genes encoding

the P. falciparum circumsporozoite (CS) protein [Dane et al.,

Science 225 : 593 (1984)], and the influenza virus non-structural

protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845

(1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1

(NS1\_81) is linked to a synthetic sequence encoding four repeat

units from the immunodominant region, which in turn is linked via

a synthetic sequence to DNA encoding Region I contg. flanking

region less the 18 AA signal region. This is linked to DNA

encoding Region II-contg. flanking region. The Pro residue sep-

arating the Asp (at the C-terminal of the linker) from the Region

I-contg. CS flanking region is an artifact of a filled-in BamHI

site; the Gly separating the Region I and II-contg. CS flanking

regions is an artifact of a synthetic FokI/TthIII I linker. The

peptide can be used in a vaccine for protection against malaria.

See also AAR12306-R12311 and AAR13175-R13179.

CC the P. falciparum circumsporozoite (CS) protein [Dame et al.,  
 CC Science 225 : 593 (1984)], and the influenza virus non-structural  
 CC protein 1 (NS1). [Baez et al., Nucleic Acids Research, 8 : 5845  
 CC (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1  
 CC (NS1-81) is linked to a synthetic sequence encoding four repeat  
 CC units (the variant form) from the immunodominant region, which in  
 CC turn is linked via a synthetic sequence to DNA encoding Region I  
 CC contg. flanking region less the 18 AA signal region. This is  
 CC linked to DNA encoding Region II-contg. flanking region. The pro  
 CC residue separating the Asp (at the C-terminal of the linker) from  
 CC in BamHI site; the Gly separating the Region I and II-contg. CS  
 CC flanking regions is an artifact of a synthetic FOKI/ThriII I  
 CC linker. The peptide can be used in a vaccine for protection  
 CC against malaria.  
 CC See also AAR12306-R12311 and AAR13175-R13178.  
 XX Sequence 335 AA;  
 SQ

Query Match 100.0%; Score 75; DB 12; Length 335;  
 Best Local Similarity 100.0%; Pred. No. 0.00043;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKLKPQGDGPN 13  
 Db 195 khkkikpgdgpn 207  
 |||||

RESULT 12  
 AAY70278  
 ID AAY70278 standard; Protein: 350 AA.  
 XX AC AAY70278;  
 XX DT 06-JUN-2000 (first entry)  
 XX DE Recombinant vaccine CDC/NIIMALVAC-1.  
 XX KW Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;  
 KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;  
 KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;  
 KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;  
 KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;  
 KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;  
 KW Pfq27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody;  
 KW honey bee.  
 XX OS Chimeric - Apis sp.  
 OS Chimeric - Clostridium tetani.  
 OS Chimeric - Plasmodium falciparum.  
 XX FH Key Location/Qualifiers  
 FT Peptide 1..22  
 FT /label= Melittin\_signal\_peptide  
 FT /note= "Derived from Honey bee"  
 FT Protein 23..350  
 FT /label= Mature\_CDC/NIIMALVAC-1  
 FT /note= "Recombinant multivalent malarial vaccine"  
 XX PN WO200011179-A1.  
 XX PD 02-MAR-2000.  
 XX PF 19-AUG-1999; 99WO-US18869.  
 XX PR 21-AUG-1998; 98US-0097703.  
 XX PA (NAIM-) NAT INST IMMUNOLOGY.  
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 XX PI Lal AA, Shi YP, Hasnain SE;  
 XX

DR WPI; 2000-237654/20.  
 DR N-PSDB; AAZ51336.  
 XX  
 PT Novel recombinant protein as vaccine for treating malarial infection  
 PT comprises antigenic peptides obtained from different stages of  
 PT plasmodium falciparum life cycle -  
 XX Claim 3; Page 43-44; 52pp; English.  
 XX  
 CC The present sequence is that of recombinant protein CDC/NIIMALVAC-1,  
 CC which is a multivalent, multistage malarial vaccine. The recombinant  
 CC protein comprises, melittin signal peptide, (His)6 tag, T-cell epitope  
 CC from tetanus toxoid and 21 antigenic epitopes from circumsporozoite  
 CC protein (CSP), sporozoite surface protein-2 (SSP-2), liver stage  
 CC antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1), MSP-2, apical  
 CC membrane antigen-1 (AMA-1), erythrocyte binding antigen-175 (EBA-175),  
 CC rhoptry associated protein-1 (RAP-1) and gamete specific antigen, Pfq27.  
 CC These epitopes were obtained at different stages of the life cycle of  
 CC Plasmodium falciparum. CDC/NIIMALVAC-1 vaccine has antiparasitic  
 CC activity and can be used for treatment and prevention of malarial  
 CC infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for detecting  
 CC P. falciparum in biological samples.  
 XX Sequence 350 AA;  
 SQ

Query Match 100.0%; Score 75; DB 21; Length 350;  
 Best Local Similarity 100.0%; Pred. No. 0.00045;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKLKPQGDGPN 13  
 Db 29 khkkikpgdgpn 41  
 |||||

RESULT 13  
 AAP83144  
 ID AAP83144 standard; protein: 411 AA.  
 XX AC AAP83144;  
 XX DT 20-NOV-1990 (first entry)  
 XX DE Sequence encoded by the circumsporozoite (CS) gene from  
 DE Plasmodium falciparum.  
 XX KW Vaccine; antigen; immunogen; probe; hybridisation;  
 KW immunoassay; diagnosis.  
 XX OS Plasmodium falciparum.  
 XX FH Key Location/Qualifiers  
 FT Region 106..120  
 FT /note= "Region 1"  
 FT Region 147..206  
 FT /note= "Repeat region, repeat unit = NANP"  
 FT Region 123..146  
 FT /note= "Repeat region, repeat unit = NANPNVDP"  
 FT Region 211..286  
 FT /note= "Repeat region, repeat unit = NANP"  
 XX PN EP278940-A.  
 XX PD 17-AUG-1988.  
 XX PF 25-JAN-1988; 88EP-0870008.  
 XX PR 30-JAN-1987; 87US-0009325.  
 XX PA (SMIK) SMITH KLINE-RIT.  
 XX PI Cabezon T, De Wilde M, Harford N;  
 XX

WPI; 1988-229751/33.  
N-PSDB; AAN81108.

DNA encoding hepatitis B virus antigens and hybrids contg. them -  
used for expression in yeast to obtain vaccines and bivalent  
vaccines

Example; Fig 3Aa-3Af; 101pp; English.

Sequence of the CS gene (AAN81108) is from lambda-mpfl. A recombinant  
DNA molecule is claimed, comprising functional DNA coding sequence fused,  
in phase, to a portion of the Pre-S2 region of a hepatitis B virus (HBV)  
Pre-S2-S protein coding sequence. The functional DNA coding sequence  
comprises the Pre-S2 coding sequence, Pre-S1 coding sequence or entire  
Pre-S1-Pre-S2 coding sequence, the CS protein coding sequence of  
Plasmodium, or a HIV coding sequence such as  
an HIV envelope gene sequence, e.g. HIV C7 protein coding region, HIV  
Peptide 121 coding region, or HIV Dreesman peptide coding region.

Sequence 411 AA;

Query Match 100.0%; Score 75; DB 9; Length 411;  
Best Local Similarity 100.0%; Pred. No. 0.00053;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHKLKLPQGNGNP 13  
|||||  
DB 108 khkkkqpqdgndp 120

RESULT 14  
AAP60416  
ID AAP60416 standard; Protein; 412 AA.  
XX AAP60416;  
AC  
XX  
DT 13-JUN-1991 (first entry)  
XX  
DE CS protein of malaria parasite.  
XX Sporozoite; vaccination.  
XX Plasmodium falciparum.  
OS  
XX Key Location/Qualifiers  
FH 124..127  
FT Region /label= Repeat unit  
FT FT  
XX EP166410-A.  
PN  
XX  
XX 02-JAN-1986.  
PD  
XX  
XX 24-JUN-1985; 85EP-0107794.  
PF  
XX  
XX 26-JUN-1984; 84US-0624564.  
PR  
XX  
XX (USDC ) US SEC OF COMMERCE.  
PA (USGO ) US GOVERNMENT.  
XX (USSA ) US SEC OF THE ARMY.  
XX  
XX McCutchan TF, Dame JB, Williams JL, Schneider I;  
PI  
XX WPI; 1986-008635/02.  
DR N-PSDB; AAN60362.  
DR

New Immunologically active pure synthetic peptide(s) - used for  
protection against infection by malaria parasite.

"  
Disclosure; Fig 2; 49pp; English.

The Plasmodium CS gene was used to isolate peptides capable of  
inducing an immune response to the parasite. Peptide antigens may

CC transformed host cell, a method of culturing the cell to produce CS, the  
CC protein, and a vaccine. The vector comprises a 192 bp Sau3A fragment  
CC coding for 16 tetrapeptide repeats of the P. falciparum CS protein  
CC derived from Sau3A digestion of a 1215 bp StuI-RsaI fragment of WR201  
CC containing the P. falciparum CS protein coding sequence minus  
CC approx. its first 50 bp, or two, three, four or more tandem copies of  
CC such 192 bp Sau3A fragment.  
XX  
SQ Sequence 412 AA;

Query Match 100.0%; Score 75; DB 9; Length 412;  
Best Local Similarity 100.0%; Pred. No. 0.00053;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KHKLLKQPGDGNP 13  
Db 109 khkllkpgdgnp 121

Search completed: January 29, 2002, 10:21:42  
Job time: 418 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:26:35 ; Search time 144.96 Seconds  
(without alignments)  
6.831 Million cell updates/sec

Title: US-09-763-397A-5

Perfect score: 75

Sequence: 1 KHKLLKQPGDGNP 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_68:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	412	1 OZ2OAF	circumsporozoite p
2	75	100.0	424	2 A54533	circumsporozoite p
3	75	100.0	442	2 A34529	circumsporozoite p
4	69	92.0	189	2 B29795	circumsporozoite p
5	69	92.0	405	2 S05428	circumsporozoite p
6	68	90.7	171	2 A29795	circumsporozoite p
7	51	68.0	388	2 A39756	circumsporozoite p
8	46	61.3	97	2 S03067	gene R protein - h
9	46	61.3	101	1 ASLJR3	vpr protein - sm1
10	43	57.3	304	2 A32993	transcription fact
11	43	57.3	944	2 T18627	hypothetical prote
12	42	56.0	179	2 B82143	conserved hypothet
13	42	56.0	978	2 C59237	pyruvate dehydroge
14	41.5	55.3	343	2 A29319	circumsporozoite p
15	41.5	55.3	367	2 A32068	circumsporozoite p
16	41.5	55.3	386	2 A48571	circumsporozoite p
17	41	54.7	102	2 A42452	circumsporozoite p
18	41	54.7	320	2 H71259	V1 protein - tobac
19	41	54.7	355	2 A42347	probable membrane
20	41	54.7	355	2 I31319	opsin, green-sensi
21	41	54.7	363	1 OZ2QAK	RH2 opsin - green
22	40	53.3	96	2 H69042	circumsporozoite p
23	40	53.3	97	2 S74326	ribosomal protein
24	40	53.3	378	1 OZ2QAB	hypothetical prote
25	40	53.3	633	2 T05883	circumsporozoite p
26	40	53.3	739	2 T21431	ATP-dependent heli
27	40	53.3	803	2 E81804	hypothetical prote
28	40	53.3	974	1 URH0AP	peptidylglycine mo
29	40	53.3	1572	2 T00376	KIAA0645 protein -

30	39.5	52.7	478	2 JQ2034	RNA-directed RNA p
31	39	52.0	101	2 S08439	vpr protein - huma
32	39	52.0	400	1 H70921	cytochrome P450 Rv
33	39	52.0	452	2 D64583	hypothetical prote
34	39	52.0	558	2 T48150	stress-induced pro
35	39	52.0	685	2 B82606	conjugal transfer
36	39	52.0	877	2 T10085	lipoygenase (EC 1
37	39	52.0	972	1 URBOAP	peptidylglycine mo
38	39	52.0	1627	2 S65464	pregnancy-associat
39	38	50.7	195	2 E70446	N-terminus of phag
40	38	50.7	278	2 E82129	rRNA (guanine-N1-)
41	38	50.7	327	2 A56553	homotetic protein H
42	38	50.7	395	2 A41156	circumsporozoite p
43	38	50.7	435	2 C84809	hypothetical prote
44	38	50.7	435	2 I38079	OXAL homolog - hum
45	38	50.7	455	2 D96525	protein T1N15.19 [

ALIGNMENTS

RESULT 1

OZZOAF

circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate ITM222)  
C:Species: Plasmodium falciparum  
C:Date: 15-Nov-1984 #sequence\_revision 15-Nov-1984 #text\_change 09-Jun-2000  
C:Accession: A03388

R:Dame, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W  
Science 225, 593-599, 1984  
A:Title: Structure of the gene encoding the immunodominant surface antigen on the spo

A:Reference number: A03388; MUID:84250215

A:Accession: A03388

A:Molecule type: DNA

A:Residues: 1-412 <DAM>

A:Cross-references: GB:K02194; NID:gl60160; PIDN:AAA29524.1; PID:gl60161

A:Experimental source: clone 7G8

C:Comment: Residues 1-16 are the probable signal sequence.

C:Comment: There are 41 copies of a 4-residue repeating unit in the middle domain of

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:336-390/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 75; DB 1; Length 412;

Best Local Similarity 100.0%; Pred. No. 4.4e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHKLLKQPGDGNP 13

|||||

Db 109 KHKLLKQPGDGNP 121

RESULT 2

A54533

circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, Thai  
C:Species: Plasmodium falciparum  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jun-2000

C:Accession: A54533

R:del Portillo, H.A.; Nussenzweig, R.S.; Enea, V.

Mol. Biochem. Parasitol. 24, 289-294, 1987

A:Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.

A:Reference number: A54533; MUID:87315205

A:Accession: A54533

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-424 <DEL>

A:Cross-references: GB:M19752; NID:gl60216; PIDN:AAA29555.1; PID:gl60217

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:348-402/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 75; DB 2; Length 424;

Best Local Similarity 100.0%; Pred. No. 4.5e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jun-2000  
 C:Accession: S05428; A45527; I60657

Qy 1 KHKLLKQPGDGNP 13  
 |||||

Db 109 KHKLLKQPGDGNP 121

## RESULT 3

A:Species: Plasmodium falciparum (strain Wellcome)  
 C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jun-2000  
 C:Accession: A54529  
 R:Lockyer, M.J.; Schwarz, R.T.  
 Mol. Biochem. Parasitol. 22, 101-108, 1987  
 A:Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum.  
 A:Reference number: A54529; MUID:87115616  
 A:Accession: A54529  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-442 <LOC>  
 A:Cross-references: GB:M15505; NID:gl60214; PIDN:AAA2955A.1; PID:gl60215  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 C:Keywords: tandem repeat  
 F:366-420/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 75; DB 2; Length 442;

Best Local Similarity 100.0%; Pred. No. 4.7e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHKLLKQPGDGNP 13

|||||

Db 119 KHKLLKQPGDGNP 131

## RESULT 4

A:Species: Plasmodium falciparum (strain LE5) (fragment)  
 C:Date: 18-Oct-1989 #sequence\_revision 21-Jul-1995 #text\_change 09-Jun-2000  
 C:Accession: B29795; A60657  
 R:de la Cruz, V.F.; Lal, A.A.; McCutchan, T.F.  
 J. Biol. Chem. 262, 11935-11939, 1987  
 A:Title: Sequence variation in putative functional domains of the circumsporozoite protein  
 A:Reference number: A92609; MUID:87308186  
 A:Accession: B29795  
 A:Molecule type: DNA  
 A:Residues: 1-161 <DEL>  
 A:Cross-references: GB:M17802; GB:M17803; GB:M17806  
 R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.  
 Mol. Biochem. Parasitol. 37, 275-280, 1989  
 A:Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell epitope  
 A:Reference number: A60657; MUID:90114334  
 A:Accession: A60657  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 152-189 <LOC>  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

Query Match 92.0%; Score 69; DB 2; Length 189;

Best Local Similarity 92.3%; Pred. No. 0.00021;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KHKLLKQPGDGNP 13

|||||

Db 92 KHKLLKQPGDGNP 104

## RESULT 5

A:Species: Plasmodium falciparum (isolate NF54)  
 C:Date: 14-Feb-1992 #sequence\_revision 14-Feb-1992 #text\_change 20-Aug-1999  
 C:Accession: A39756  
 R:Lal, A.A.; Goldman, I.F.

C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jun-2000  
 C:Accession: S05428; A45527; I60657  
 R:Campbell, J.R.  
 Nucleic Acids Res. 17, 5854, 1989  
 A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate  
 A:Reference number: S05428; MUID:893645189  
 A:Accession: S05428  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-405 <CAM>  
 A:Cross-references: EMBL:X15363  
 R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.  
 Mol. Biochem. Parasitol. 37, 275-280, 1989  
 A:Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate  
 A:Reference number: A45527; MUID:89364998  
 A:Accession: A45527  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-405 <CAS>  
 A:Cross-references: GB:M22982; GB:J04650; NID:gl60168; PIDN:AAA29527.1; PID:gl60169  
 R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.  
 Mol. Biochem. Parasitol. 37, 275-280, 1989  
 A:Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell  
 A:Reference number: A60657; MUID:90114334  
 A:Accession: I60657  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 319-336,354-373 <LOC>  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 F:329-383/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 92.0%; Score 69; DB 2; Length 405;

Best Local Similarity 92.3%; Pred. No. 0.00045;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KHKLLKQPGDGNP 13

|||||

Db 90 KHKLLKQPGDGNP 102

## RESULT 6

A:Species: Plasmodium falciparum (strain B11) (fragment)  
 C:Date: 28-Aug-1989 #sequence\_revision 28-Aug-1989 #text\_change 09-Jun-2000  
 C:Accession: A29795  
 R:de la Cruz, V.F.; Lal, A.A.; McCutchan, T.F.  
 J. Biol. Chem. 262, 11935-11939, 1987  
 A:Title: Sequence variation in putative functional domains of the circumsporozoite protein  
 A:Reference number: A92609; MUID:87308186  
 A:Accession: A29795  
 A:Molecule type: DNA  
 A:Residues: 1-171 <DELA>  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

Query Match 90.7%; Score 68; DB 2; Length 171;

Best Local Similarity 92.3%; Pred. No. 0.00028;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KHKLLKQPGDGNP 13

|||||

Db 102 KHKLLKQPGDGNP 114

## RESULT 7

A:Species: Plasmodium reichenowi  
 C:Date: 14-Feb-1992 #sequence\_revision 14-Feb-1992 #text\_change 20-Aug-1999  
 C:Accession: A39756  
 R:Lal, A.A.; Goldman, I.F.



J. Biol. Chem. 266, 6686-6689, 1991  
A:Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria  
A:Reference number: A39756; MUID:91201303  
A:Accession: A39756  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-388 <LAL>  
A:Cross-references: GB:M60972; NID:g160228; PIDN:AAA29561.1; PID:g160229  
C:Superfamily: Circumsporozoite protein; thrombospondin type 1 repeat homology  
F:312-366/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 68.0%; Score 51; DB 2; Length 388;  
Best Local Similarity 75.0%; Pred. No. 0.51;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KHKLKQPGDGN 12  
I : : ||| ||| :  
Db 109 KHKLKQPGDGN 120

RESULT 8  
S03067  
gene R protein - human T-cell lymphotropic virus type 4  
C:Species: human T-cell lymphotropic virus type 4, HTLV-4  
C:Date: 21-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 20-Sep-1999  
C:Accession: S03067  
R:Hahn, B.H.; Kong, L.I.; Lee, S.W.; Kumar, P.; Taylor, M.E.; Arya, S.K.; Shaw, G.M.  
Nature 300, 184-186, 1987  
A:Title: Relation of HTLV-4 to simian and human immunodeficiency-associated viruses.  
A:Reference number: S03065  
A:Accession: S03067  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-97 <HAH>  
A:Cross-references: EMBL:X06391; NID:g61580; PIDN:CAA29689.1; PID:g61582  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1987  
C:Superfamily: AIDS vpr protein

Query Match 61.3%; Score 46; DB 2; Length 97;  
Best Local Similarity 58.3%; Pred. No. 0.91;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HKKLKQPGDGNP 13  
I : : ||| |||  
Db 79 HSRIGQPGGPN 90

RESULT 9  
ASLJR3  
vpr protein - simian immunodeficiency virus (macaque isolate)  
N:Alternate names: orf-R protein  
C:Species: simian immunodeficiency virus, SIV  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 16-Jul-1999  
C:Accession: E28887  
R:Chakrabarti, L.; Guyader, M.; Allizon, M.; Daniel, M.D.; Desrosiers, R.C.; Tiollais, P.  
Nature 328, 543-547, 1987  
A:Title: Sequence of simian immunodeficiency virus from macaque and its relationship to  
A:Reference number: A28887; MUID:87287230  
A:Accession: E28887  
A:Molecule type: DNA  
A:Residues: 1-101 <CHA>  
A:Cross-references: GB:Y00277; GB:M16403; NID:g61730; PIDN:CAA68383.1; PID:g61735  
C:Genetics:  
A:Gene: vpr  
C:Superfamily: AIDS vpr protein  
C:Keywords: AIDS; immunodeficiency

Query Match 61.3%; Score 46; DB 1; Length 101;  
Best Local Similarity 58.3%; Pred. No. 0.95;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HKKLKQPGDGNP 13  
I : : ||| |||  
Db 79 HSRIGQPGGPN 90

RESULT 10  
A32993  
transcription factor GATA-1 - chicken  
N:Alternate names: transcription factor Eryf1  
C:Species: Gallus gallus (chicken)  
C:Date: 29-Dec-1989 #sequence\_revision 29-Dec-1989 #text\_change 26-Aug-1999  
C:Accession: A32993  
R:Evans, T.; Felsenfeld, G.  
Cell 58, 877-885, 1989  
A:Title: The erythroid-specific transcription factor Eryf1: a new finger protein.  
A:Reference number: A32993; MUID:89376538  
A:Accession: A32993  
A:Molecule type: mRNA  
A:Residues: 1-304 <EVA>  
A:Cross-references: GB:M26209; NID:g212628; PIDN:AAA49055.1; PID:g212629  
C:Superfamily: transcription factor GATA-1; GATA-type zinc finger homology  
C:Keywords: DNA binding; nucleus; transcription factor; transcription regulation; zinc  
F:107-160/Domain: GATA-type zinc finger homology <GzF1>  
F:110-134/Region: zinc finger GATA motif  
F:161-214/Domain: GATA-type zinc finger homology <GzF2>  
F:164-188/Region: zinc finger GATA motif

Query Match 57.3%; Score 43; DB 2; Length 304;  
Best Local Similarity 61.5%; Pred. No. 9.3;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KHKLKQPGDGNP 13  
I : : ||| |||  
Db 218 KKKRRPPGGNP 230

RESULT 11  
TI8627  
hypothetical protein B0001.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: TI8627  
R:Sims, M.  
submitted to the EMBL Data Library, February 1996  
A:Reference number: Z18999  
A:Accession: TI8627  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-944 <WIL>  
A:Cross-references: EMBL:Z69634; PIDN:CAA93451.1; GSPDB:GN00022; CESP:B0001.2  
A:Experimental source: Clone B0001  
C:Genetics:  
A:Gene: CESP:B0001.2  
A:Map position: 4  
A:Introns: 18/2; 75/3; 119/2; 350/3; 475/2; 487/3; 521/2; 601/3; 641/3; 749/2; 870/1

Query Match 57.3%; Score 43; DB 2; Length 944;  
Best Local Similarity 72.7%; Pred. No. 29;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KHKLKQPGDG 11  
I : : ||| |||  
Db 309 KHKLFQAGDG 319

RESULT 12  
B82143  
conserved hypothetical protein VC1892 [imported] - Vibrio cholerae (strain N16961 ser  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: B82143  
 R:Heidelberq, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 chardson, D.; Ermlaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 A:Reference number: A82035; MUID:20406833  
 A:Accession: B82143  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-179 <HEI>  
 A:Cross-references: GB:AE004265; GB:AE003852; NID:9956424; PIDN:AAF95040.1; GSPDB:GN001  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VC1892  
 A:Map position: 1  
 C:Superfamily: *Escherichia coli* ycfp protein

Query Match 56.0%; Score 42; DB 2; Length 179;  
 Best Local Similarity 58.3%; Pred. No. 8.1;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Qy 2 HKKLKQPGDGNP 13  
 ||::|||  
 Db 51 HKAIEQSGDNP 62  
 RESULT 13  
 C59237  
 pyruvate dehydrogenase (EC 1.2.-.-) dihydrolipoyl transacetylase/lipoamide dehydrogenase  
 C:Species: *Thiobacillus ferrooxidans*  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000  
 C:Accession: C59237  
 R:Powles, R.; Deane, S.; Rawlings, D.  
 Microbiology 142, 2543-2548, 1996  
 A:Title: The gene for gamma-glutamylcysteine synthetase from *Thiobacillus ferrooxidans*  
 A:Reference number: A59237; MUID:96425880  
 A:Accession: C59237  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-978 <POW>  
 A:Cross-references: GB:U81808; GB:L75931; NID:gl750277; PIDN:AAB41628.1; PID:gl750280  
 A:Experimental source: strain ATCC 33020  
 C:Genetics:  
 A:Gene: pdhc  
 C:Keywords: oxidoreductase

Query Match 56.0%; Score 42; DB 2; Length 978;  
 Best Local Similarity 70.0%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Qy 2 HKKLKQPGDG 11  
 ||::|||  
 Db 301 HKKQPGDG 310  
 RESULT 14  
 A29319  
 circumsporozoite protein - *Plasmodium vivax* (strain Sal-I) (fragment)  
 C:Species: *Plasmodium vivax*  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 17-Mar-1999  
 C:Accession: A29319; A44187  
 R:de la Cruz, V.F.; Lal, A.A.; Welsh, J.A.; McCutchan, T.F.  
 J. Biol. Chem. 262, 6464-6467, 1987  
 A:Title: Evolution of the immunodominant domain of the circumsporozoite protein gene from  
 A:Reference number: A29319; MUID:87194878  
 A:Accession: A29319  
 A:Molecule type: DNA  
 A:Residues: 1-343 <DEL>  
 R:McCutchan, T.F.; Lal, A.A.; de la Cruz, V.F.; Miller, L.H.; Maloy, W.L.; Charoenvit, X  
 Science 230, 1381-1383, 1985

A:Title: Sequence of the immunodominant epitope for the surface protein on sporozoite  
 A:Reference number: A44187; MUID:86070222  
 A:Accession: A44187  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-98 <MCC>  
 A:Cross-references: GB:J02751  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 F:268-321/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match 55.3%; Score 41.5; DB 2; Length 343;  
 Best Local Similarity 56.2%; Pred. No. 19;  
 Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;  
 Qy 1 KHKKLKQPG---DGNP 13  
 : ||||| ||  
 Db 54 RENKLKQPGDRADGQP 69

RESULT 15  
 A32068  
 circumsporozoite protein - *Plasmodium vivax* (strain North Korean) (fragment)  
 C:Species: *Plasmodium vivax*  
 C:Date: 21-May-1990 #sequence\_revision 21-May-1990 #text\_change 20-Aug-1999  
 C:Accession: A32068  
 R:Arnot, D.E.; Barnwell, J.W.; Stewart, M.J.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 8102-8106, 1988  
 A:Title: Does biased gene conversion influence polymorphism in the circumsporozoite p  
 A:Reference number: A32068; MUID:89042133  
 A:Accession: A32068  
 A:Molecule type: DNA  
 A:Residues: 1-367 <ARN>  
 A:Cross-references: GB:M20670; GB:J04090; NID:gl60183; PIDN:AAA29534.1; PID:g552194  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 F:292-345/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match 55.3%; Score 41.5; DB 2; Length 367;  
 Best Local Similarity 56.2%; Pred. No. 20;  
 Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;  
 Qy 1 KHKKLKQPG---DGNP 13  
 : ||||| ||  
 Db 62 RENKLKQPGDRADGQP 77

Search completed: January 29, 2002, 10:26:36  
 Job time: 647 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:24:02 ; Search time 133.18 seconds  
(without alignments)  
2.197 Million cell updates/sec

Title: US-09-763-397A-5  
Perfect score: 75  
Sequence: 1 KHKLLKQPGDGNP 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTOUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	22	6 5178861-5	Patent No. 5178861
2	75	100.0	412	1 US-08-313-288B-18	Sequence 18, Appl
3	56	74.7	13	6 5178861-17	Patent No. 5178861
4	43	57.3	135	3 US-08-937-271-22	Sequence 22, Appl
5	41.5	55.3	232	6 5171843-7	Patent No. 5171843
6	41.5	55.3	378	6 5171843-9	Patent No. 5171843
7	41	54.7	27	6 5178861-4	Patent No. 5178861
8	40	53.3	1020	2 US-08-070-301-3	Sequence 3, Appl
9	39	52.0	989	2 US-08-070-301-14	Sequence 14, Appl
10	38	50.7	10	2 US-08-318-856A-15	Sequence 15, Appl
11	37	49.3	9	2 US-08-318-856A-14	Sequence 14, Appl
12	37	49.3	442	2 US-08-821-355A-5	Sequence 5, Appl
13	37	49.3	442	2 US-09-003-687A-5	Sequence 5, Appl
14	37	49.3	442	4 US-09-136-605-5	Sequence 5, Appl
15	37	49.3	521	1 US-08-075-193-2	Sequence 2, Appl
16	37	49.3	521	5 PCT-US94-06698-2	Sequence 2, Appl
17	37	49.3	521	5 PCT-US94-06698-2	Sequence 2, Appl
18	37	49.3	596	2 US-08-821-355A-6	Sequence 6, Appl
19	37	49.3	596	2 US-09-003-687A-6	Sequence 6, Appl
20	37	49.3	596	4 US-09-136-605-6	Sequence 6, Appl
21	36	48.0	907	4 US-08-989-299-12	Sequence 12, Appl
22	36	48.0	1964	2 US-08-790-912-3	Sequence 3, Appl
23	36	48.0	2052	2 US-08-790-912-2	Sequence 2, Appl
24	35	48.7	26	1 US-07-942-245-446	Sequence 446, Appl
25	35	46.7	50	1 US-08-363-311-21	Sequence 21, Appl
26	35	46.7	50	2 US-08-463-288A-21	Sequence 21, Appl
27	35	46.7	50	2 US-08-470-445A-21	Sequence 21, Appl

28	35	46.7	50	2	US-08-462-679-21	Sequence 21, Appl
29	35	46.7	50	2	US-08-466-210A-21	Sequence 21, Appl
30	35	46.7	50	2	US-08-467-147A-21	Sequence 21, Appl
31	35	46.7	50	2	US-08-469-014-21	Sequence 21, Appl
32	35	46.7	50	5	PCT-US93-10506A-21	Sequence 21, Appl
33	35	46.7	50	5	PCT-US93-10506-21	Sequence 21, Appl
34	35	46.7	737	1	US-08-188-582-16	Sequence 16, Appl
35	35	46.7	737	1	US-08-646-715-16	Sequence 16, Appl
36	35	46.7	848	1	US-08-045-806-4	Sequence 4, Appl
37	35	46.7	848	1	US-08-366-051B-4	Sequence 4, Appl
38	35	46.7	2414	1	US-08-227-536-2	Sequence 2, Appl
39	35	46.7	2414	1	PCT-US95-04682-2	Sequence 2, Appl
40	34	45.3	15	6	5171843-11	Patent No. 5171843
41	34	45.3	148	1	US-08-450-257-59	Sequence 59, Appl
42	34	45.3	148	1	US-08-450-246-59	Sequence 59, Appl
43	34	45.3	148	1	US-08-450-098-59	Sequence 59, Appl
44	34	45.3	148	1	US-08-451-233-59	Sequence 59, Appl
45	34	45.3	148	1	US-08-450-236-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1  
5178861-5  
; Patent No. 5178861  
; APPLICANT: VERGARA, ULISES; RUIZ, ANDRES; FERREIRA, ARTURO;  
; NUSSENZWEIG, RUTH S.; NUSSENZWEIG, VICTOR N.  
; TITLE OF INVENTION: CROSS-REACTIVE AND PROTECTIVE EPITOPES  
; OF CIRCUMPOROZOITE PROTEINS  
; NUMBER OF SEQUENCES: 18  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370,241  
; FILING DATE: 22-JUN-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 115,634  
; FILING DATE: 26-OCT-1987  
; APPLICATION NUMBER: 649,903  
; FILING DATE: 12-SEP-1984  
; SEQ ID NO: 5:  
; LENGTH: 22  
5178861-5

Query Match 100.0%; Score 75; DB 6; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHKLLKQPGDGNP 13  
|||||  
DB 5 KHKLLKQPGDGNP 17

RESULT 2  
US-08-313-288B-18  
; Sequence 18, Application US/08313288B  
; Patent No. 5750502  
; GENERAL INFORMATION:  
; APPLICANT: Jessell, Thomas M. and Avi Hu Klar  
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/313,288B  
; FILING DATE: January 5, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0526  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 412 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-313-288B-18

Query Match 100.0%; Score 75; DB 1; Length 412;  
Best Local Similarity 100.0%; Pred. No. 0.0004;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHKLLKQPGDGNP 13  
Db 109 KHKLLKQPGDGNP 121

RESULT 3  
5178861-17  
; Patent No. 5178861  
; APPLICANT: VERGARA, ULISES; RUIZ, ANDRES; FERREIRA, ARTURO;  
; NUSSENZWEIG, RUTH S.; NUSSENZWEIG, VICTOR N.  
; TITLE OF INVENTION: CROSS-REACTIVE AND PROTECTIVE EPITOPES  
; OF CIRCUMPOROZOITE PROTEINS  
; NUMBER OF SEQUENCES: 18  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370,241  
; FILING DATE: 22-JUN-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 115,634  
; FILING DATE: 26-OCT-1987  
; APPLICATION NUMBER: 649,903  
; FILING DATE: 12-SEP-1984  
; SEQ ID NO: 17:  
; LENGTH: 13  
5178861-17

Query Match 74.7%; Score 56; DB 6; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHKLLKQPGD 10  
Db 4 KHKLLKQPGD 13

RESULT 4  
US-08-937-271-22  
; Sequence 22, Application US/08937271  
; Patent No. 6063386  
; GENERAL INFORMATION:  
; APPLICANT: Dale, James B.  
; APPLICANT: Lederer, James W.  
; TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN  
; TITLE OF INVENTION: VACCINE  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/937,271  
; FILING DATE: 15-SEP-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rosenman, Stephen J.  
; REGISTRATION NUMBER: 43,058  
; REFERENCE/DOCKET NUMBER: 481112.405C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 135 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-937-271-22

Query Match 57.3%; Score 43; DB 3; Length 135;  
Best Local Similarity 61.5%; Pred. No. 9.9;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KHKLLKQPGDGNP 13  
Db 80 RHVLLDNGDGNP 92

RESULT 5  
5171843-7  
; Patent No. 5171843  
; APPLICANT: NUSSENZWEIG, VICTOR  
; TITLE OF INVENTION: IMMUNOGENIC POLYPEPTIDE AND METHOD FOR  
; PURIFYING IT  
; NUMBER OF SEQUENCES: 13  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/175,112  
; FILING DATE: 30-MAR-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 754,645  
; FILING DATE: 9-JUL-1985  
; APPLICATION NUMBER: 115,634  
; FILING DATE: 26-OCT-1987  
; APPLICATION NUMBER: 649,903  
; FILING DATE: 12-SEP-1984  
; SEQ ID NO: 7:  
; LENGTH: 232  
5171843-7

Query Match 55.3%; Score 41.5; DB 6; Length 232;  
Best Local Similarity 56.2%; Pred. No. 28;  
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

Qy 1 KHKLLKQPG---DGNP 13  
Db 7 RENKLLKQPGDRADGQP 22

RESULT 6  
5171843-9

Patent No. 5171843  
; APPLICANT: NUSSENZWEIG, VICTOR  
; TITLE OF INVENTION: IMMUNOGENIC POLYPEPTIDE AND METHOD FOR  
; PURIFYING IT  
; NUMBER OF SEQUENCES: 13  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/175.112  
; FILING DATE: 30-MAR-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 754.645  
; FILING DATE: 9-JUL-1985  
; APPLICATION NUMBER: 115.634  
; FILING DATE: 26-OCT-1987  
; APPLICATION NUMBER: 649.903  
; FILING DATE: 12-SEP-1984  
; SEQ ID NO:9:  
; LENGTH: 378  
5171843-9

Query Match 55.3%; Score 41.5; DB 6; Length 378;  
Best Local Similarity 56.2%; Pred. No. 45;  
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

Oy 1 KHKLKQPG---DGNP 13  
; : ||||| || |  
Db 89 RENLKQPGDRADGQ 104

RESULT 7  
5178861-4  
; Patent No. 5178861  
; APPLICANT: VERGARA, ULISES; RUIZ, ANDRES; FERREIRA, ARTURO;  
; NUSSENZWEIG, RUTH S.; NUSSENZWEIG, VICTOR N.  
; TITLE OF INVENTION: CROSS-REACTIVE AND PROTECTIVE EPITOPES  
; OF CIRCUMFOROZOITE PROTEINS  
; NUMBER OF SEQUENCES: 18  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370.241  
; FILING DATE: 22-JUN-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 115.634  
; FILING DATE: 26-OCT-1987  
; APPLICATION NUMBER: 649.903  
; FILING DATE: 12-SEP-1984  
; SEQ ID NO:4:  
; LENGTH: 27  
5178861-4

Query Match 54.7%; Score 41; DB 6; Length 27;  
Best Local Similarity 70.0%; Pred. No. 4.2;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 4 KLKQPGDGNP 13  
; ||||| : |  
Db 8 KLKQPNEGOP 17

RESULT 8  
US-08-070-301-3  
; Sequence 3, Application US/08070301  
; Patent No. 5871995  
; GENERAL INFORMATION:  
; APPLICANT: IIDA, Toshio  
; APPLICANT: KAMINUMA, Toshihiko  
; APPLICANT: FUSE, Yuka  
; APPLICANT: TAJIMA, Masahiro  
; APPLICANT: YANAGI, Mitsuo  
; APPLICANT: OKAMOTO, Hiroshi  
; APPLICANT: KISHIMOTO, Jiro  
; APPLICANT: IFUKU, Ohji  
; APPLICANT: KATO, Ichiro

; TITLE OF INVENTION: ENZYME PARTICIPATING IN C-TERMINAL  
; TITLE OF INVENTION: AMIDATION, AND METHOD OF PREPARING SAME AND USE THEREOF  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wegner, Cantor, Mueller & Player, P.C.  
; STREET: 1233 20th Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20036-8218  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/070.301  
; FILING DATE: 24-MAY-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 1-209687  
; FILING DATE: 15-AUG-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 1-181933  
; FILING DATE: 31-OCT-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-76331  
; FILING DATE: 26-MAR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-106412  
; FILING DATE: 24-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-205475  
; FILING DATE: 02-AUG-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Player, William E.  
; REGISTRATION NUMBER: 31,409  
; REFERENCE/DOCKET NUMBER: P-450-22830  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-040  
; TELEFAX: (202) 835-0605  
; TELEX: 440706  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1020 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-070-301-3

Query Match 53.3%; Score 40; DB 2; Length 1020;  
Best Local Similarity 53.8%; Pred. No. 2e+02;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 1 KHKLKQPGDGNP 13  
; : ||||| |  
Db 899 KQKLIKEPGSGVP 911

RESULT 9  
US-08-070-301-14  
; Sequence 14, Application US/08070301  
; Patent No. 5871995  
; GENERAL INFORMATION:  
; APPLICANT: IIDA, Toshio  
; APPLICANT: KAMINUMA, Toshihiko  
; APPLICANT: FUSE, Yuka  
; APPLICANT: TAJIMA, Masahiro  
; APPLICANT: YANAGI, Mitsuo  
; APPLICANT: OKAMOTO, Hiroshi  
; APPLICANT: KISHIMOTO, Jiro  
; APPLICANT: IFUKU, Ohji

APPLICANT: KATO, Ichiro  
TITLE OF INVENTION: ENZYME PARTICIPATING IN C-TERMINAL  
AMIDATION, AND METHOD OF PREPARING SAME AND USE THEREOF  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wegner, Cantor, Mueller & Player, P.C.  
STREET: 1233 20th Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20036-8218  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/070,301  
FILING DATE: 24-MAY-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-209687  
FILING DATE: 15-AUG-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-181933  
FILING DATE: 31-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-76331  
FILING DATE: 26-MAR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-106412  
FILING DATE: 24-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-205475  
FILING DATE: 02-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,409  
REFERENCE/DOCKET NUMBER: P-450-22830  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-040  
TELEFAX: (202) 835-0605  
TELEX: 440706  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 989 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Bovine  
US-08-070-301-14

Query Match 52.0%; Score 39; DB 2; Length 989;  
Best Local Similarity 53.8%; Pred. No. 2.7e+02;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KHKLLKQPGDGNP 13  
| :|:|:| :|  
DB 867 KQKLVKPGSGVP 879

RESULT 10  
US-08-318-856A-15  
; Sequence 15, Application US/08318856A  
; Patent No. 5972351  
; GENERAL INFORMATION:  
; APPLICANT: Adrian V. S. Hill, et al.  
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-  
; RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE  
; TITLE OF INVENTION: ANTIGENS (AS AMENDED)

NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
STREET: 2033 K Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1+  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,856A  
FILING DATE: October 3, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 92 08 068.8  
FILING DATE: April 3, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 92 17 704.7  
FILING DATE: August 20, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/GB93/00711  
FILING DATE: April 5, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee Cheng  
REGISTRATION NUMBER: 40,949  
REFERENCE/DOCKET NUMBER: 263-PFIR15770US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 721-8200  
TELEFAX: (202) 721-8250  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acid residues  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-318-856A-15

Query Match 50.7%; Score 38; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.7;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PGDGNP 13  
|:|:|:|  
DB 2 PGDGNP 7

RESULT 11  
US-08-318-856A-14  
; Sequence 14, Application US/08318856A  
; Patent No. 5972351  
; GENERAL INFORMATION:  
; APPLICANT: Adrian V.S. Hill, et al.  
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-  
; RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE  
; TITLE OF INVENTION: ANTIGENS (AS AMENDED)  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 5.1+



```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318.856A
; FILING DATE: October 3, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 08 068.8
; FILING DATE: April 3, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 17 704.7
; FILING DATE: August 20, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/00711
; FILING DATE: April 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 263-PP1R1577US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 721-8200
; TELEFAX: (202) 721-8250
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-318-856A-14

```

```

Query Match          49.3%; Score 37; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 KHKKLKQ 7
Db 3 KHKKLKQ 9

```

```

RESULT 12
US-08-821-355A-5
; Sequence 5, Application US/08821355A
; Patent No. 5851775
; GENERAL INFORMATION:
; APPLICANT: Barker, Nick
; APPLICANT: Clevers, Hans
; APPLICANT: Korinek, Vladimir
; APPLICANT: Morin, Patrice
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Vogelstein, Bert
; APPLICANT: Sparks, Andrew
; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
; TITLE OF INVENTION: Interact to Prevent Cancer
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821.355A
; FILING DATE: 20-MAR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32,145
; REFERENCE/DOCKET NUMBER: 1107.05064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 97430 BMB UT
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 442 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5851775e
US-08-821-355A-5

```

```

Query Match          49.3%; Score 37; DB 2; Length 442;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 KHKKLKQPCDGN 12
Db 405 KKRDRKQPGETN 416

```

```

RESULT 13
US-09-003-687A-5
; Sequence 5, Application US/09003687A
; Patent No. 5998600
; GENERAL INFORMATION:
; APPLICANT: Barker, Nick
; APPLICANT: Clevers, Hans
; APPLICANT: Korinek, Vladimir
; APPLICANT: Morin, Patrice
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Vogelstein, Bert
; APPLICANT: Sparks, Andrew
; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
; TITLE OF INVENTION: Interact to Prevent Cancer
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/003.687A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/821.355
; FILING DATE: 20-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32,145
; REFERENCE/DOCKET NUMBER: 1107.05064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 97430 BMB UT
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 442 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5998600e
US-09-003-687A-5

Query Match 49.3%; Score 37; DB 2; Length 442;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KHKLKQPGDGN 12
Db 405 KKRDKQPGETN 416

RESULT 14
US-09-136-605-5
; Sequence 5, Application US/09136605A
; Patent No. 6140052
; GENERAL INFORMATION:
; APPLICANT: He, Tong-Chuan
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC Interact to
; TITLE OF INVENTION: Prevent Cancer
; FILE REFERENCE: 1107.75741
; CURRENT APPLICATION NUMBER: US/09/136,605A
; CURRENT FILING DATE: 1998-08-20
; EARLIER APPLICATION NUMBER: 08/821,355
; EARLIER FILING DATE: 1997-03-20
; EARLIER APPLICATION NUMBER: 09/003,687
; EARLIER FILING DATE: 1998-01-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-136-605-5

Query Match 49.3%; Score 37; DB 4; Length 442;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KHKLKQPGDGN 12
Db 405 KKRDKQPGETN 416

RESULT 15
US-08-075-193-2
; Sequence 2, Application US/08075193
; Patent No. 5547868
; GENERAL INFORMATION:
; APPLICANT: MILLER, WALTER L.
; APPLICANT: HARIKRISHNA, JENNIFER A.
; APPLICANT: BLACK, STEPHEN M.
; TITLE OF INVENTION: CHOLESTEROL DISPOSAL FUSION ENZYMES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/075,193
```

```
;
; FILING DATE: 09-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY Ph.D., RICHARD L.
; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: UCAL-236/000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-494-7622
; TELEFAX: 415-857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-075-193-2

Query Match 49.3%; Score 37; DB 1; Length 521;
Best Local Similarity 45.5%; Pred. No. 3e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 HKKLKQPGDGN 12
Db 183 HRRIKKAGSGN 193

Search completed: January 29, 2002, 10:24:02
Job time: 508 sec
```



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:13:38 ; Search time 80.65 seconds  
(without alignments)  
5.910 Million cell updates/sec

Title: US-09-763-397A-5  
Perfect score: 75  
Sequence: 1 KHKKLKQPCDGNP 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues  
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	412	1	CSP_PLAFA
2	75	100.0	424	1	CSP_PLAFA
3	75	100.0	442	1	CSP_PLAFA
4	69	92.0	315	1	CSP_PLAFL
5	69	92.0	397	1	CSP_PLAFL
6	51	68.0	388	1	CSP_PLARE
7	47	62.7	102	1	GCCI_HUMAN
8	46	61.3	97	1	VPR_SIVMK
9	46	61.3	101	1	VPR_SIVM1
10	43	57.3	304	1	GATL_CHICK
11	41.5	55.3	343	1	CSP_PLAYS
12	41.5	55.3	378	1	CSP_PLAYS
13	41.5	55.3	386	1	CSP_PLAST
14	41	54.7	102	1	YLIK_TYDVA
15	41	54.7	355	1	OPSB_ANOCA
16	41	54.7	355	1	OPSB_CHICK
17	41	54.7	363	1	CSP_PLAKH
18	40	53.3	96	1	RI21_MERTH
19	40	53.3	101	1	VPR_SIVSP
20	40	53.3	378	1	CSP_PLACB
21	40	53.3	974	1	AMD_HUMAN
22	40	53.3	4036	1	RRPL_DUGBV
23	39	52.0	101	1	VPR_HV2D2
24	39	52.0	400	1	CI41_MYCTU
25	39	52.0	972	1	AMD_BOVIN
26	38	50.7	327	1	HXD1_MOUSE
27	38	50.7	495	1	OXA1_HUMAN
28	38	50.7	520	1	C11A_BOVIN
29	38	50.7	520	1	C11A_SHEEP
30	38	50.7	1912	1	CHD4_HUMAN
31	37	49.3	80	1	AKA7_MOUSE
32	37	49.3	349	1	OPSH_CARAU
33	37	49.3	349	1	OPSH_CARAU

34	37	49.3	354	1	OPSD_RAUER
35	37	49.3	354	1	OPSD_SCYCA
36	37	49.3	355	1	OPSB_GECCO
37	37	49.3	521	1	C11A_HUMAN
38	37	49.3	907	1	MOP_DESGT
39	37	49.3	979	1	AMD_MOUSE
40	36	48.0	45	1	VG78_BPMU5
41	36	48.0	186	1	25SE_BRANA
42	36	48.0	318	1	MPD1_YEAST
43	36	48.0	354	1	OPSL_ASTFA
44	36	48.0	361	1	BRN4_HUMAN
45	36	48.0	361	1	BRN4_MOUSE

ALIGNMENTS

RESULT 1  
ID CSP\_PLAFA STANDARD: PRT: 412 AA.  
AC P02893:  
BT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).  
OS Plasmodium falciparum.  
CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OK NCBI\_TaxID=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84250215; PubMed=6204383;  
RA Dane J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
RA Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
RA Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;  
RT "Structure of the gene encoding the immunodominant surface antigen on  
the sporozoite of the human malaria parasite Plasmodium falciparum.";  
RL Science 225:593-599(1984).  
CC -I- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT  
SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE  
MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE  
VERTEBRATE HOST).  
CC -I- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR  
ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES  
WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.  
CC -I- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
CC EMBL: K02194; AAA29524.1; -  
DR PIR: A03388; OZ20AF.  
DR InterPro: IPR003067; Crcmsprzotite.  
DR InterPro: IPR000884; TSP1.  
DR Pfam: PF00090; TSP\_1; 1.  
DR PRINTS: PR01303; Crcmsprzotite.  
DR SMART: SM00209; TSP1; 1.  
KW Malaria; Sporozoite; Repeat; Signal.  
FT SIGNAL 1 16 PROBABLY.  
FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.  
FT DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.  
SO SEQUENCE 412 AA; 44420 MM; 1EEED3DE905F8 CRC64;  
  
Query Match 100.0%; Score 75; DB 1; Length 412;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KHKKLKQPCDGNP 13

```

Db 109 KHKRLKOPGDGNP 121

|||||
RESULT 2
CSP_PLAFT STANDARD; PRT; 424 AA.
AC P13814;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium falciparum (isolate t4 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;

RN
SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373.
RA del Portillo H.A., Nussenzweig R.S., Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from
   Thailand."
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19752; AAA2955.1; -.
DR PIR; A54533; A54533.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PRO1303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 424 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;

Query Match 100.0%; Score 75; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 1,4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHKRLKOPGDGNP 13
|||||
Db 109 KHKRLKOPGDGNP 121

```

```

RN
SEQUENCE FROM N.A.
RX MEDLINE=87115616; PubMed=3543671;
RA Lockyer M.J., Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium
   falciparum."
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15505; AAA2955.4.1; -.
DR PIR; A54529; A54529.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PRO1303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 442 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 130 320 47 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;

Query Match 100.0%; Score 75; DB 1; Length 442;
Best Local Similarity 100.0%; Pred. No. 1,4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHKRLKOPGDGNP 13
|||||
Db 119 KHKRLKOPGDGNP 131

RESULT 4
CSP_PLAFT STANDARD; PRT; 315 AA.
ID CSP_PLAFT
AC P05691;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (CS) (FRAGMENT).
OS Plasmodium falciparum (isolate le5).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5840;

RN
SEQUENCE FROM N.A.
RX MEDLINE=87308186; PubMed=2442154;
RA la Cruz V.F., Lal A.A., McCutchan T.F.;
RT "Sequence variation in putative functional domains of the
   circumsporozoite protein of Plasmodium falciparum. Implications for
   vaccine development."
RL J. Biol. Chem. 262:11935-11939(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.

```

```

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17802; AAA29538.1; -
KW Malaria; Sporozoit; Repeat.
FT DOMAIN 1 1 40 X 4 AA TANDEM REPEATS OF P-N-A-N.
FT NON_TER 107 270 315
FT TER 315 315
SQ SEQUENCE 315 AA; 33649 MW; A334DB11FA7FD777 CRC64;

Query Match
Best Local Similarity 92.0%; Score 69; DB 1; Length 315;
Matches 12; Conservative 0; Pred. No. 0.00011; Mismatches 1; Indels 0; Gaps 0;

OY 1 KHKKLKOPDGNP 13
Db 93 KHKKLKOPADGNP 105

RESULT 5
CSP_PLAFO
ID CSP_PLAFO STANDARD: PRT: 397 AA.
AC P19597; Q25798;
DT 01-FEB-1991 (Rel. 17, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345189; Pubmed=2668895;
RA Campbell J.R.;
RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria
RT candidate vaccine antigen."
RL Nucleic Acids Res. 17:5854-5854 (1989).
RN [2]
RP REVISIONS.
RA Campbell J.R.;
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92155298; Pubmed=1346766;
RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,
RA Thomas A.W., Bagar S., Cochran M.A., Thanassi J., Levine M.M.,
RA Hackett C.S.;
RT "Plasmodium falciparum: in vitro characterization and human
RT infectivity of a cloned line."
RL Exp. Parasitol. 74:159-168 (1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89364998; Pubmed=2671723;
RA Caspers P., Gentz R., Matile H., Pink J.R., Sinaigalla F.;
RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum
RT isolate used in malaria vaccine trials."
RL Mol. Biochem. Parasitol. 35:185-190 (1989).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15363; CAA33421.1; -
DR EMBL: M83886; AAA29521.1; -
DR EMBL: M22982; AAA29527.1; -
DR PIR: S05428; S05428.
DR PIR: A45527; A45527.
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSPL.
DR Pfam: PF00090; TSP_1.1.
DR PRINTS: PR01303; CIRCMSPRZOITE.
DR SMART: SM00209; TSPL.1.
KW Malaria; Sporozoit; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT CONFLICT 194 194 A -> ANPANDNA (IN REF. 4).
SQ SEQUENCE 397 AA; 42646 MW; 9EB1146F59BCEA3 CRC64;

Query Match
Best Local Similarity 92.0%; Score 69; DB 1; Length 397;
Matches 12; Conservative 0; Pred. No. 0.00014; Mismatches 1; Indels 0; Gaps 0;

OY 1 KHKKLKOPDGNP 13
Db 90 KHKKLKOPADGNP 102

RESULT 6
CSP_PLARE
ID CSP_PLARE STANDARD: PRT: 388 AA.
AC P26594;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91201303; Pubmed=2016283;
RA Lal A.A., Goldman I.F.;
RT "Circumsporozoite protein gene from Plasmodium reichenowi, a
RT chimpanzee malaria parasite evolutionarily related to the human
RT malaria parasite Plasmodium falciparum."
RL J. Biol. Chem. 266:6686-6689 (1991).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60972; AAA29561.1; -
DR PIR: A39756; A39756.

```

```

DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000084; TSPL.
DR Pfam: PF00090; tsp.1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSPL; 1.
FT SIGNAL 1 16 PROBABILE.
FT CHAIN 17 388 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 388 AA; 42245 MW; C031EFBE2E35604 CRC64;

Query Match
Best Local Similarity 68.0%; Score 51; DB 1; Length 388;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HKHKLKOPGDCN 12
DB 109 HKHKLKOPGDCN 120

RESULT 7
GGCL_HUMAN STANDARD; PRT; 102 AA.
AC 060829;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE G ANTIGEN FAMILY C.1 PROTEIN (PROSTATE-ASSOCIATED GENE PROTEIN 4)
DE (PAGE-4) (PAGE-1) (OM27) (GAGE-9).
GN GAGEC1 OR PAGE4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RA MEDLINE=98393718; PubMed=9724777;
RA Brinkmann U., Vasmataz G., Lee B., Yerushalmi N., Essand M.,
RA "Pastan I., an X chromosome-linked GAGE-like gene that is expressed in
RT normal and neoplastic prostate, testis, and uterus."
RL Proc. Natl. Acad. Sci. U.S.A. 95:10757-10762(1998).
RN [2]
RN SEQUENCE FROM N.A.
RA Strom T.M., Nyakatura G., Hellebrand H., Drescher B., Rosenthal A.,
RA Meindl A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: UNKNOWN.
CC -I- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN NORMAL MALE AND
CC FEMALE REPRODUCTIVE TISSUES, PROSTATE, TESTIS, FALLOPIAN TUBE,
CC UTERUS, AND PLACENTA, AS WELL AS IN PROSTATE CANCER, TESTICULAR
CC CANCER, AND UTERINE CANCER.
CC -I- SIMILARITY: BELONGS TO THE GAGE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL; AF275258; AAF88037.1; -
CC EMBL; AJ005894; CAA06751.1; -
CC EMBL; AF238380; AAF62541.1; -
CC MIM; 300287; -
SQ SEQUENCE 102 AA; 11153 MW; CE5D07AFBF73301B CRC64;

Query Match
Best Local Similarity 62.7%; Score 47; DB 1; Length 102;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 HKHKLKOPGDCNP 13
DB 90 HKHKLKOPGDCNP 102

RESULT 8
VPR_SIVMK STANDARD; PRT; 97 AA.
ID VPR_SIVMK
AC P05957;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DE VPR PROTEIN (R ORF PROTEIN).
GN VPR.
OS Simian immunodeficiency virus (K6W isolate) (SIV-MAC).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11735;
[1]
RN SEQUENCE FROM N.A.
RA MEDLINE=87287229; PubMed=3497350;
RA Franchini G., Gallo R.C., Guo H.-G., Gurgo C., Callatti E.,
RA Fargnoli K., Hall L., Wong-Straal F., Reitz M.S. Jr.;
RT "Sequence of simian immunodeficiency virus and its relationship to
RL the human immunodeficiency viruses."
RL Nature 328:539-543(1987).
CC -I- MISCELLANEOUS: THIS IS A MACAQUE ISOLATE.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL; M19499; AAB59909.1; ALT_SEQ.
DR HIV; M15897; VPRSK6W78.
DR InterPro: IPR000012; HIV_ORFPR.
DR Pfam: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVPRVPX.
KW AIDS.
SQ SEQUENCE 97 AA; 11005 MW; FD7932FC6CE5362 CRC64;

Query Match
Best Local Similarity 61.3%; Score 46; DB 1; Length 97;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 HKHKLKOPGDCNP 13
DB 79 HSRIGOPGDCNP 90

RESULT 9
VPR_SIVM1 STANDARD; PRT; 101 AA.
ID VPR_SIVM1
AC P05958;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DE VPR PROTEIN (R ORF PROTEIN).
GN VPR.
OS Simian immunodeficiency virus (Mm142-83 isolate) (SIV-MAC).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11733;
[1]
RN SEQUENCE FROM N.A.
RA MEDLINE=87287230; PubMed=3649576;
RA Chakrabarti L., Guyader M., Allison M., Daniel M.D., Desrosiers R.C.,
RA Tiollais P., Sonigo P.;
RT "Sequence of simian immunodeficiency virus from macaque and its
RT relationship to other human and simian retroviruses."

```



```

RL Nature 328:543-547(1987).
CC -I- MISCELLANEOUS: THIS IS A MACAQUE ISOLATE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y00277; CA66383.1; -.
DR PIR; E28887; ASLJR3.
DR HIV; M16403; VPR5MM142.
DR InterPro: IPR000012; HIV_ORFXR.
DR Pfam: PF00522; VPR.1.
DR PRINTS; PR00444; HIVPRVFX.
KM AIDS.
SQ SEQUENCE 101 AA; 11461 MW; A1F62078F596C0B9 CRC64;

Query Match          61.3%; Score 46; DB 1; Length 101;
Best Local Similarity 58.3%; Pred. NO. 0.32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 HKKLKOPGDNP 13
   1 1111111
Db 79 HSRIGQPGGNNP 90

RESULT 10
GATL_CHICK STANDARD: PRT; 304 AA.
ID GATL_CHICK
AC P17678;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ERYTHROID TRANSCRIPTION FACTOR (GATA-1) (ERYF1) (NF-E1 DNA-BINDING
DE PROTEIN) (NF-E1A).
GN GAT1 OR ERYF1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=8937638; PubMed=2776214;
RT Evans T., Felsenfeld G.;
RT "The erythroid-specific transcription factor Eryf1: a new finger
RT protein.";
RL Cell 58:877-885(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=91065513; PubMed=2249770;
RT Yamamoto M., Ko L.J., Leonardo M.W., Beug H., Orkin S.H.;
RA Engel J.D.;
RT "Activity and tissue-specific expression of the transcription factor
RT NF-E1 multigene family.";
RL Genes Dev. 4:1650-1662(1990).
RN [3]
RP STRUCTURE BY NMR OF 158-223.
RA MEDLINE=93324913; PubMed=8332909;
RT Omichinski J.G., Clore G.M., Schaad O., Felsenfeld G., Traitor C.,
RA Appella E., Stahl S.J., Gronenborn A.M.;
RT "NMR structure of a specific DNA complex of Zn-containing DNA binding
RT domain of GATA-1.";
RL Science 261:438-446(1993).
RN [4]
RP STRUCTURE BY NMR OF 158-223.
RA MEDLINE=9744676; PubMed=9303001;
RT Tjandra N., Omichinski J.G., Gronenborn A.M., Clore G.M., Bax A.;
RT "Use of dipolar 1H-15N and 1H-13C couplings in the structure

```

```

RT determination of magnetically oriented macromolecules in solution.";
RL Nat. Struct. Biol. 4:732-738(1997).
CC -I- FUNCTION: TRANSCRIPTIONAL ACTIVATOR WHICH PROBABLY SERVES AS A
CC GENERAL SWITCH FACTOR FOR ERYTHROID DEVELOPMENT. IT BINDS TO DNA
CC SITES WITH THE CONSENSUS SEQUENCE (A/T)GATA(A/G) WITHIN REGULATORY
CC REGIONS OF GLOBIN GENES AND OF OTHER GENES EXPRESSED IN ERYTHROID
CC CELLS.
CC -I- SUBCELLULAR LOCATION: NUCLEAR.
CC -I- TISSUE SPECIFICITY: ERYTHROCYTE.
CC -I- DOMAIN: THE TWO FINGERS ARE FUNCTIONALLY DISTINCT AND COOPERATE TO
CC ACHIEVE SPECIFIC, STABLE DNA BINDING. THE FIRST FINGER IS
CC NECESSARY ONLY FOR FULL SPECIFICITY AND STABILITY OF BINDING,
CC WHEREAS THE SECOND ONE IS REQUIRED FOR BINDING (BY SIMILARITY).
CC -I- SIMILARITY: CONTAINS 2 GATA-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26209; AAA49055.1; -.
DR PIR; A32993; A32993.
DR PDB; 1GAT; 31-OCT-93.
DR PDB; 2GAT; 28-JAN-98.
DR PDB; 3GAT; 28-JAN-98.
DR PDB; 1GAU; 31-OCT-93.
DR TRANSFAC; T00267; -.
DR InterPro: IPR002965; P-rich_extensn.
DR InterPro: IPR000679; ZnF_GATA.
DR InterPro: IPR001164; ZnF_GCS.
DR Pfam; PF00320; GATA.2.
DR PRINTS; PR00619; GATAZNFINGER.
DR SMART; SM00401; ZnF_GATA.2.
DR PROSITE; PS00344; GATA_ZN_FINGER_1; 2.
DR PROSITE; PS0114; GATA_ZN_FINGER_2; 2.
KM Zinc-finger; Zinc-finger; Activator; DNA-binding; Erythrocyte;
KM Zinc-finger; Nuclear protein; 3D structure.
FT ZN_FING 110 134 GATA-TYPE 1.
FT ZN_FING 110 134 GATA-TYPE 2.
FT ZN_FING 164 188
FT TURN 160 161
FT TURN 165 167
FT STRAND 175 178
FT TURN 179 181
FT STRAND 182 185
FT HELIX 186 195
FT TURN 196 196
FT TURN 202 203
SQ SEQUENCE 304 AA; 31417 MW; 64C9D6FDB58CE83F CRC64;

Query Match          57.3%; Score 43; DB 1; Length 304;
Best Local Similarity 61.5%; Pred. NO. 3.5;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 HKKLKOPGDNP 13
   1 1111111
Db 218 KKKRRRPPGNNP 230

RESULT 11
CSP_PLAIVS STANDARD: PRT; 343 AA.
ID CSP_PLAIVS
AC P13826;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCUMSPORZOITE PROTEIN (CS) (FRAGMENT).
OS Plasmodium vivax (strain Salvador I).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5856;

```

```

RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=66070222; PubMed=2416057;
RA McCutchan T.F., Lal A.A., de la Cruz V.F., Miller L.H., Maloy W.L.,
RA Chareonvilt Y., Beaudoin R.L., Guerry P., Wister R.J., Hoffman S.L.,
RA Hockmeyer W.T., Collins W.E., Wirth D.;
RT "Sequence of the immunodominant epitope for the surface protein on
RT sporozoites of Plasmodium vivax.";
RL Science 230:1381-1383(1985).
RN [2]
RX SEQUENCE FROM N.A.
RA MEDLINE=67194878; PubMed=2437120;
RA de la Cruz V.F., Lal A.A., Welsh J.A., McCutchan T.F.;
RT "Evolution of the immunodominant domain of the circumsporozoite
RT protein gene from Plasmodium vivax. Implications for vaccines.";
RL J. Biol. Chem. 262:6464-6467(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; TSP_1; 1.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
KM Sporozoite; Malaria; Repeat.
FT NON_TER 1
FT DOMAIN 63 243 19 X 9 AA TANDEM REPEATS OF P-G-D-R-A-D-
FT SEQUENCE 343 AA; 34155 MW; 308EPD5BHC15DFC3 CRC64;

Query Match 55.38; Score 41.5; DB 1; Length 343;
Best Local Similarity 56.28; Pred. No. 7.2;
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 KHKRLKQPG--DGNP 13
Db 54 RKNKLRQPGDRADGQ 69
: ||||| | |

```

```

CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL: M11926; AAA29526.1; -.
DR EMBL: J02751; AAA29529.1; ALT_SEQ.
DR PIR: A26256; OZ2OAV.
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; TSP_1; 1.
DR PRINTS: PR01303; CIRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
KM Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1
FT CHAIN 20 378
FT DOMAIN 97 267
FT FT 19 CIRCUMSPOROZOITE PROTEIN.
FT FT 19 X 9 AA TANDEM REPEATS OF P-G-D-R-A-D-
FT FT G-O-P.
FT FT 36 36 G->E (IN REF. 1).
FT FT 96 96 G->R (IN REF. 1).
FT FT 295 295 E->A (IN REF. 1).
FT FT 328 328 R->S (IN REF. 1).
FT FT CONFLICT 328 328
FT SEQUENCE 378 AA; 37800 MW; C84B5BED05E3CED CRC64;

Query Match 55.38; Score 41.5; DB 1; Length 378;
Best Local Similarity 56.28; Pred. No. 8;
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 KHKRLKQPG--DGNP 13
Db 88 RKNKLRQPGDRADGQ 103
: ||||| | |

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L05068; AAA29525.1; -  
DR PIR; A48571; A48571.  
DR InterPro: IPR003067; Circmsprzote.  
DR InterPro: IPR000884; TSP1.  
DR Pfam; PF00090; TSP\_1; 1.  
DR PRINTS; PR01303; CIRCMSPRZOTE.  
DR SMART; SM00209; TSP1; 1.  
DR PROSITE; PS50092; TSP1; 1.  
KW Malaria; Sporozoite; Signal; Repeat.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 386 CIRCUMSPOROZOITE PROTEIN.  
FT DOMAIN 96 275 20 X 9 AA TANDEM REPEATS OF G-D-R-A-[AD]-  
FT G-O-P-A.  
SQ SEQUENCE 386 AA; 38567 MW; A0097D4BDE5548DB CRC64;

Query Match 55.3%; Score 41.5; DB 1; Length 386;  
Best Local Similarity 56.2%; Pred. No. 8.2;  
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

OY 1 KHKKIKQPC---DGNP 13  
: ||||| |||  
DB 88 RENKIKQPCDRAQGP 103

RESULT 14  
Y1K\_TYDVA  
ID Y1K\_TYDVA STANDARD; PRT; 102 AA.  
AC P31619;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-OCT-1993 (Rel. 27, Last annotation update)  
DE HYPOTHETICAL 11.2 KDA PROTEIN.  
CN  
OS Tobacco yellow dwarf virus (strain Australia) (TYDV).  
OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.  
OX NCBI\_TaxID=31599;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92188538; Pubmed=1546458;  
RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;  
RT "The nucleotide sequence of the infectious cloned DNA component of  
RT tobacco yellow dwarf virus reveals features of geminiviruses  
RT infecting monocotyledonous plants.";  
RL Virology 187:633-642(1992).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M81103; AAA47947.1; -  
DR PIR; A42452; A42452.  
DR InterPro: IPR002621; Geminlmov.  
DR Pfam; PF01708; Geminl\_mov; 1.  
DR Prodom; PD004500; Geminl\_mov; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 102 AA; 11178 MW; AA0CEFL0AF55B67 CRC64;

Query Match 54.7%; Score 41; DB 1; Length 102;  
Best Local Similarity 66.7%; Pred. No. 2.4;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 LKQPGDGNP 13  
: ||||: |||  
DB 89 LRRPGECP 97

RESULT 15  
OPSB\_ANOCA  
ID OPSB\_ANOCA STANDARD; PRT; 355 AA.  
AC P51471;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE BLUE-SENSITIVE OPSIN (BLUE PHOTORECEPTOR PIGMENT) (RH2 OPSIN).  
OS Anolis carolinensis (Green anole) (American chameleon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Iguania; Iguanidae; Polychrotinae; Anolis.  
OX NCBI\_TaxID=28377;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95371134; Pubmed=7643409;  
RA Kawamura S., Yokoyama S.; rhodopsinlike opsin genes in lizards.";  
RT "Paralogous origin of the rhodopsinlike opsin genes in lizards.";  
RL J. Mol. Evol. 40:594-600(1995).  
CC -I- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT  
CC MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY  
CC LINKED TO CIS-RETINAL. THIS OPSIN USES A VITAMIN-A2 CHROMOPHORE.  
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -I- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY  
CC BE PHOSPHORYLATED (BY SIMILARITY).  
CC -I- MISCELLANEOUS: THIS OPSIN HAS PROBABLY AN ABSORPTION MAXIMUM AT  
CC 503 NM.  
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC OPSIN SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; S79167; AAB35062.1; -  
DR EMBL; S79124; AAB35062.1; JOINED.  
DR EMBL; S79134; AAB35062.1; JOINED.  
DR EMBL; S79165; AAB35062.1; JOINED.  
DR EMBL; S79166; AAB35062.1; JOINED.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR InterPro: IPR001760; Opsin.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCR\_RHODPSN.  
DR PRINTS; PR00238; OPSIN.  
DR PRINTS; PR00579; RHODOPSIN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
DR PROSITE; PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
DR PROSITE; PS00238; OPSIN; 1.  
KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;  
KW Phosphorylation; G-protein coupled receptor; Lipoprotein; Palmitate.  
FT DOMAIN 1 36  
FT TRANSMEM 37 61  
FT DOMAIN 62 73  
FT TRANSMEM 74 98  
FT DOMAIN 99 113  
FT TRANSMEM 114 133  
FT DOMAIN 134 152  
FT TRANSMEM 153 176  
FT TRANSMEM 177 202  
FT TRANSMEM 203 230  
FT TRANSMEM 231 252  
FT TRANSMEM 253 276  
FT DOMAIN 277 284  
FT TRANSMEM 285 309

FT	DOMAIN	310	355	CYTOPLASMIC.
FT	CARBOHYD	2	2	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	15	15	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	200	200	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	110	187	BY SIMILARITY.
FT	BINDING	296	296	RETINAL CHROMOPHORE (BY SIMILARITY).
FT	LIPID	322	322	PALMITATE (BY SIMILARITY).
FT	LIPID	323	323	PALMITATE (BY SIMILARITY).
SO	SEQUENCE	355 AA;	39983 MW;	555B671A4DE96EA6 CRC64;

Query Match 54.7%; Score 41; DB 1; Length 355;  
 Best Local Similarity 87.5%; Pred. NO. 9.1;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKLKOP 8  
 |||||:  
 Db 64 KKKLKOP 71

Search completed: January 29, 2002, 11:13:38  
 Job time: 814 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:12:09 ; Search time 285.36 Seconds  
(without alignments)  
6.664 Million cell updates/sec

Title: US-09-763-397A-5

Perfect score: 75

Sequence: 1 KHKLLKQPGDGNP 13

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_17:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	75	100.0	383	5 Q9GPN1	Q9gpn1 plasmodium
2	75	100.0	416	5 Q25829	Q25829 plasmodium
3	75	100.0	420	5 Q25831	Q25831 plasmodium
4	75	100.0	420	5 Q25838	Q25838 plasmodium
5	75	100.0	424	5 Q27425	Q27425 plasmodium
6	75	100.0	424	5 Q92556	Q92556 plasmodium
7	75	100.0	432	5 Q27246	Q27246 plasmodium
8	75	100.0	432	5 Q25827	Q25827 plasmodium
9	75	100.0	436	5 Q27325	Q27325 plasmodium
10	75	100.0	436	5 Q25828	Q25828 plasmodium
11	75	100.0	442	5 Q25830	Q25830 plasmodium
12	75	100.0	452	5 Q25834	Q25834 plasmodium
13	69	92.0	393	5 Q92555	Q92555 plasmodium
14	69	92.0	408	5 Q25729	Q25729 plasmodium
15	48	64.0	101	12 Q88146	Q88146 chimpanzee
16	47	62.7	101	12 Q76637	Q76637 human immun
17	47	62.7	126	12 Q74231	Q74231 human immun
18	46	61.3	92	12 Q88042	Q88042 chimpanzee
19	46	61.3	92	12 Q88059	Q88059 chimpanzee

20	46	61.3	92	12	Q88021	Q88021 chimpanzee
21	46	61.3	92	12	Q88041	Q88041 chimpanzee
22	46	61.3	92	12	Q88043	Q88043 chimpanzee
23	46	61.3	92	12	Q88045	Q88045 chimpanzee
24	46	61.3	92	12	Q88047	Q88047 chimpanzee
25	46	61.3	92	12	Q88048	Q88048 chimpanzee
26	46	61.3	92	12	Q88051	Q88051 chimpanzee
27	46	61.3	92	12	Q88053	Q88053 chimpanzee
28	46	61.3	92	12	Q88055	Q88055 chimpanzee
29	46	61.3	92	12	Q88060	Q88060 chimpanzee
30	46	61.3	92	12	Q88062	Q88062 chimpanzee
31	46	61.3	92	12	Q88065	Q88065 chimpanzee
32	46	61.3	92	12	Q88066	Q88066 chimpanzee
33	46	61.3	92	12	Q88069	Q88069 chimpanzee
34	46	61.3	92	12	Q89782	Q89782 chimpanzee
35	46	61.3	97	12	Q85605	Q85605 human t-cel
36	46	61.3	97	12	Q11785	Q11785 chimpanzee
37	46	61.3	101	12	Q07392	Q07392 chimpanzee
38	46	61.3	101	12	Q88017	Q88017 chimpanzee
39	45	60.0	116	5	Q9XZE7	Q9XZE7 sterkiella
40	43	57.3	944	5	Q17411	Q17411 caenorhabdi
41	42.5	56.7	230	5	Q9BPK7	Q9bpk7 plasmodium
42	42.5	56.7	230	5	Q9BPK5	Q9bpk5 plasmodium
43	42	56.0	140	5	Q25212	Q25212 junonia coe
44	42	56.0	179	2	Q9KQV6	Q9kgv6 vibrio chol
45	42	56.0	389	13	Q13025	Q13025 xenopus lae

#### ALIGNMENTS

RESULT 1

Q9GPN1 PRELIMINARY; PRT; 383 AA.

AC Q9GPN1; DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=5833;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FCC-1/HN;

RA Zheng C., Xie P., Chen Y.;

RT "Molecular cloning and expression of circumsporozoite protein gene

from Plasmodium falciparum FCC-1/HN strain in mycobacterium.";

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF315469; AAG37074.1; .

DR InterPro; IPR003067; Crcmsprzoite.

DR Pfam; PF00090; tsp\_1; 1.

DR PRINTS; PR01303; Crcmsprzoite.

FT NON\_TER 1

FT NON\_TER 383

SQ SEQUENCE 383 AA; 40893 MW; 503C5DFDF61A9E27 CRC64;

Query Match 100.0%; Score 75; DB 5; Length 383;

Best Local Similarity 100.0%; Pred. No. 4.4e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHKLLKQPGDGNP 13

|||||

Db 90 KHKLLKQPGDGNP 102

RESULT 2

Q25829

ID Q25829 PRELIMINARY; PRT; 416 AA.

AC Q25829;

DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN.  
GN CSP.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=MAD20;  
RC MEDLINE=84250215; PubMed=6204383;  
RX Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;  
RT "Structure of the gene encoding the immunodominant surface antigen on  
RT the sporozoite of the human malaria parasite Plasmodium falciparum."  
RL Science 225:593-599(1984).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=MAD20;  
RC STRAIN=MAD20;  
RX la Cruz V.F.;  
RA Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A.  
RP STRAIN=MAD20;  
RC Jongwutiwes S., Tanabe K., Kanbara H.;  
RL Mol. Biochem. Parasitol. 0:0-0(0).  
DR EMBL: M83172; AAA29550.1; -;  
DR InterPro: IPR000884; TSP1.  
DR InterPro: IPR003067; Crmsprzoite.  
DR Pfam: PF00090; tsp\_1; 1.  
DR PRINTS; PR01303; CRMSPRZOITE.  
DR SMART; SM00209; TSP1; 1.  
DR SEQUENCE 416 AA; 44829 MW; D3EF560B2D368DE9 CRC64;  
SQ  
Query Match 100.0%; Score 75; DB 5; Length 416;  
Best Local Similarity 100.0%; Pred. No. 4.8e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KHKLLKQPGDGNP 13  
Db 109 KHKLLKQPGDGNP 121  
RESULT 3  
Q25831 PRELIMINARY; PRT; 420 AA.  
ID Q25831;  
AC Q25831;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN.  
GN CSP.  
OS Plasmodium falciparum (isolate K1 / Thailand).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=K1;  
RC MEDLINE=84250215; PubMed=6204383;  
RX Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;  
RT "Structure of the gene encoding the immunodominant surface antigen on  
RT the sporozoite of the human malaria parasite Plasmodium falciparum."  
RL Science 225:593-599(1984).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=K1;  
RC la Cruz V.F.;  
RA Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]

RP SEQUENCE FROM N.A.  
RC STRAIN=K1;  
RA Jongwutiwes S., Tanabe K., Kanbara H.;  
RL Mol. Biochem. Parasitol. 0:0-0(0).  
DR EMBL: M83174; AAA29552.1; -;  
DR InterPro: IPR000884; TSP1.  
DR InterPro: IPR003067; Crmsprzoite.  
DR Pfam: PF00090; tsp\_1; 1.  
DR PRINTS; PR01303; CRMSPRZOITE.  
DR PROSITE; PS50092; TSP1; 1.  
DR SMART; SM00209; TSP1; 1.  
DR SEQUENCE 420 AA; 45318 MW; F7F70F1C4939DEA7 CRC64;  
SQ  
Query Match 100.0%; Score 75; DB 5; Length 420;  
Best Local Similarity 100.0%; Pred. No. 4.8e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KHKLLKQPGDGNP 13  
Db 109 KHKLLKQPGDGNP 121  
RESULT 4  
Q25838 PRELIMINARY; PRT; 420 AA.  
ID Q25838;  
AC Q25838;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN.  
GN CSP.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=835B;  
RC MEDLINE=84250215; PubMed=6204383;  
RX Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;  
RT "Structure of the gene encoding the immunodominant surface antigen on  
RT the sporozoite of the human malaria parasite Plasmodium falciparum."  
RL Science 225:593-599(1984).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=835B;  
RC la Cruz V.F.;  
RA Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A.  
RP STRAIN=835B;  
RC Jongwutiwes S., Tanabe K., Kanbara H.;  
RL Mol. Biochem. Parasitol. 0:0-0(0).  
DR EMBL: M83161; AAA29574.1; -;  
DR InterPro: IPR000884; TSP1.  
DR InterPro: IPR003067; Crmsprzoite.  
DR Pfam: PF00090; tsp\_1; 1.  
DR PRINTS; PR01303; CRMSPRZOITE.  
DR SMART; SM00209; TSP1; 1.  
DR SEQUENCE 420 AA; 45155 MW; 3A85B92432C2893C CRC64;  
SQ  
Query Match 100.0%; Score 75; DB 5; Length 420;  
Best Local Similarity 100.0%; Pred. No. 4.8e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KHKLLKQPGDGNP 13  
Db 109 KHKLLKQPGDGNP 121

```
RESULT 5
Q27425 ID Q27425 PRELIMINARY; PRT; 424 AA.
AC Q27425:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata, Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=844;
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=844;
RA la Cruz V.F.;
RN [3]
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=844;
RA Jongcuttiwies S., Tanabe K., Kanbara H.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M83169; AAA29547.1; -
DR EMBL; M83149; AAA29562.1; -
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
SQ SEQUENCE 424 AA; 45592 MW; F20CEB60636DB98E CRC64;

Query Match 100.0%; Score 75; DB 5; Length 424;
Best Local Similarity 100.0%; Pred. No. 4.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHKLLKQPGDGNP 13
Db 109 KHKLLKQPGDGNP 121

RESULT 6
Q99256 ID Q99256 PRELIMINARY; PRT; 424 AA.
AC Q99256;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (CS) PRECURSOR, VARIANT 2 (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91270295; PubMed=2052038;
RA Lockyer M.J.;
RT "Clonal variation in the Plasmodium falciparum circumsporozoite
RT protein gene.";
RL Mol. Biochem. Parasitol. 45:179-182(1991).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
```

```
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
DR EMBL; M57499; AAA63422.1; -
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
KW Malaria; Repeat; Sporozoite; Signal.
FT SIGNAL 1 16 BY SIMILARITY.
FT CHAIN 17 >424 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 302 4-RESIDUE TANDEM REPEATS.
FT NON_TER 424 424
SQ SEQUENCE 424 AA; 45609 MW; BF6D0F906C664B8F CRC64;

Query Match 100.0%; Score 75; DB 5; Length 424;
Best Local Similarity 100.0%; Pred. No. 4.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHKLLKQPGDGNP 13
Db 109 KHKLLKQPGDGNP 121

RESULT 7
Q27246 ID Q27246 PRELIMINARY; PRT; 432 AA.
AC Q27246;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=946;
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=946;
RA la Cruz V.F.;
RN [3]
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=946;
RA Jongcuttiwies S., Tanabe K., Kanbara H.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M83155; AAA29568.1; -
DR EMBL; M83170; AAA29548.1; -
DR EMBL; M83152; AAA29565.1; -
DR EMBL; M83158; AAA29571.1; -
DR EMBL; M83166; AAA29544.1; -
DR EMBL; M83168; AAA29546.1; -
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
SQ SEQUENCE 432 AA; 46414 MW; 8787E6005578873A CRC64;

Query Match 100.0%; Score 75; DB 5; Length 432;
```

MEDLINE=84250215; PubMed=6204383;  
Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;  
"Structure of the gene encoding the immunodominant surface antigen on  
the sporozoite of the human malaria parasite Plasmodium falciparum.";  
Science 225:593-599(1984).  
[2]  
RN  
SEQUENCE FROM N.A.  
RC  
STRAIN=837;  
la Cruz V.F.;  
RRL  
Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
[3]  
RN  
SEQUENCE FROM N.A.  
RC  
STRAIN=837;  
Jongwattiwes S., Tanabe K., Kanbara H.;  
RRL  
Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RRL  
EMBL; M83164; AAA29542.1; -  
DR  
EMBL; M83150; AAA29563.1; -  
DR  
EMBL; M83163; AAA29576.1; -  
DR  
InterPro; IPR000884; TSP1.  
DR  
InterPro; IPR003067; Crmsprzoite.  
DR  
Pfam; PF00090; tsp\_1; 1.  
DR  
PRINTS; PRO1303; CRMSPRZOITE.  
DR  
SMART; SM00209; TSP1; 1.  
DR  
SEQUENCE 436 AA; 4668 MW; 5B42FF3348B68655 CRC64;  
SQ

Query Match 100.0%; Score 75; DB 5; Length 436;  
Best Local Similarity 100.0%; Pred. No. 5e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 1 KHKKKQPGDGNP 13  
|||||  
DB 109 KHKKKQPGDGNP 121

RESULT 10  
Q25828 PRELIMINARY; PRT; 436 AA.  
ID Q25828;  
AC Q25828;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN.  
GN CSP.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP  
RC  
STRAIN=842;  
RX MEDLINE=84250215; PubMed=6204383;  
DAME J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;  
"Structure of the gene encoding the immunodominant surface antigen on  
the sporozoite of the human malaria parasite Plasmodium falciparum.";  
Science 225:593-599(1984).  
[2]  
RN  
SEQUENCE FROM N.A.  
RC  
STRAIN=842;  
Jongwattiwes S., Tanabe K., Kanbara H.;  
RRL  
Mol. Biochem. Parasitol. 0:0-0(0).  
RRL  
EMBL; M83167; AAA29545.1; -  
DR  
InterPro; IPR000884; TSP1.  
DR  
InterPro; IPR003067; Crmsprzoite.  
DR  
Pfam; PF00090; tsp\_1; 1.  
DR





DR SMART; SM00209; TSP1; 1.  
KW Malaria; Repeat; Sporozoite; Signal.  
FT SIGNAL 1 16  
FT CHAIN 17 >393  
FT DOMAIN 104 267  
FT NON\_TER 393 393  
SQ SEQUENCE 393 AA; 42263 MW; 45169AE773689037 CRC64;

Query Match 92.0%; Score 69; DB 5; Length 393;  
Best Local Similarity 92.3%; Pred. No. 0.0005;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKLKQPGDGNP 13  
| | | | | | | | | |  
Db 90 KKKLKQPADGNP 102

RESULT 14  
Q25729 PRELIMINARY; PRT; 408 AA.  
AC Q25729;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE CIRCUMSPOROZOITE PROTEIN.  
GN CS.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SANTA LUCIA, SAL1;  
RA Qari S.H., Lal A.A.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U20969; AAA63153.1; -  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR003067; Crmsprzoite.  
DR Pfam; PF00090; tsp\_1; 1.  
DR PRINTS; PR01303; CRMSPRZOITE.  
DR SMART; SM00209; TSP1; 1.  
SQ SEQUENCE 408 AA; 43871 MW; A545BE517822515C CRC64;

Query Match 92.0%; Score 69; DB 5; Length 408;  
Best Local Similarity 92.3%; Pred. No. 0.00052;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKLKQPGDGNP 13  
| | | | | | | | | |  
Db 109 KKKLKQPADGNP 121

RESULT 15  
Q88146 PRELIMINARY; PRT; 101 AA.  
AC Q88146;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE VPR PROTEIN.  
GN VPR.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92124755; PubMed=173112;  
RA Novembre F.J., Hirsch V.M., McClure H.M., Fultz P.N., Johnson P.R.;  
RT "SIV from stump-tailed macaques: molecular characterization of a  
highly transmissible primate lentivirus.";  
RL Virology 186:783-787(1992).  
DR EMBL; M83293; AAA91943.1; -.

DR InterPro; IPR000012; HIV\_OREXR.  
DR Pfam; PF00522; VPR; 1.  
DR PRINTS; PR00444; HIVVPRVPX.  
SQ SEQUENCE 101 AA; 11448 MW; 2E48F2E2AA187911 CRC64;

Query Match 64.0%; Score 48; DB 12; Length 101;  
Best Local Similarity 53.8%; Pred. No. 0.62;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKKLKQPGDGNP 13  
: | : | | | | |  
Db 78 RHRIGQPGGPNP 90

Search completed: January 29, 2002, 11:12:10  
Job time: 766 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:06:40 ; Search time 353.23 Seconds  
(without alignments)  
17.293 Million cell updates/sec

Title: US-09-763-397A-25

Perfect score: 108

Sequence: 1 MKFLVNLVFMVVIYSIYAD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*

2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*

3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*

4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*

5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*

6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*

7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*

8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*

9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*

10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*

11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*

12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*

13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*

14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*

15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*

16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*

17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*

18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*

19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*

20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*

21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*

22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*

23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*

24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	108	100.0	22	21	US-09-763-397A-25	Sequence 25, Appl
2	108	100.0	129	24	US-60-292-638-6	Sequence 6, Appl
3	108	100.0	350	21	US-09-763-397A-2	Sequence 2, Appl
4	108	100.0	1189	13	US-08-911-312-34	Sequence 34, Appl
5	108	100.0	1189	13	US-08-911-312A-34	Sequence 34, Appl
6	108	100.0	1189	13	US-08-912-951-325	Sequence 325, App
7	108	100.0	1189	13	US-08-974-549-613	Sequence 613, App
8	108	100.0	1189	18	US-09-432-503-613	Sequence 4, Appl
9	102	94.4	21	17	US-09-376-330-4	

10	102	94.4	21	21	US-09-766-378-36	Sequence 36, Appl
11	102	94.4	21	21	US-09-766-378A-36	Sequence 36, Appl
12	102	94.4	70	22	US-09-847-208-26	Sequence 26, Appl
13	102	94.4	177	1	PCT-US00-02740-5	Sequence 5, Appl
14	102	94.4	177	22	US-09-890-806-5	Sequence 5, Appl
15	93.5	86.6	26	11	US-08-782-481-43	Sequence 43, Appl
16	93.5	86.6	26	11	US-08-782-481-43	Sequence 43, Appl
17	93.5	86.6	26	14	US-09-005-167-43	Sequence 43, Appl
18	93.5	86.6	26	14	US-09-005-167A-43	Sequence 43, Appl
19	93.5	86.6	26	14	US-09-005-318A-43	Sequence 43, Appl
20	93.5	86.6	26	14	US-09-005-318C-43	Sequence 43, Appl
21	93.5	86.6	26	15	US-09-176-741-43	Sequence 43, Appl
22	93.5	86.6	26	15	US-09-176-741B-43	Sequence 43, Appl
23	93.5	86.6	26	15	US-09-337-827-23	Sequence 23, Appl
24	55	50.9	205	17	US-09-553-662-23	Sequence 23, Appl
25	55	50.9	205	17	US-09-553-662-24	Sequence 24, Appl
26	54	50.0	205	17	US-09-553-662-24	Sequence 24, Appl
27	54	50.0	205	19	US-09-553-662-26	Sequence 26, Appl
28	52	48.1	305	1	PCT-US99-03055-11	Sequence 11, Appl
29	49	45.4	397	11	US-08-792-498A-6	Sequence 6, Appl
30	49	45.4	668	24	US-60-173-464-23762	Sequence 23762, A
31	49	45.4	669	24	US-60-131-637-30145	Sequence 30145, A
32	49	45.4	669	24	US-60-191-681-24136	Sequence 24136, A
33	49	45.4	73	15	US-09-134-001C-3037	Sequence 3037, Ap
34	48	44.4	73	18	US-09-450-969-5866	Sequence 5866, Ap
35	48	44.4	65	20	US-09-611-529-5585	Sequence 5585, Ap
36	47	43.5	77	18	US-09-417-507-41740	Sequence 41740, A
37	47	43.5	147	24	US-60-196-713-2480	Sequence 2480, Ap
38	47	43.5	206	8	US-08-487-032A-755	Sequence 755, App
39	47	43.5	206	9	US-08-561-469A-755	Sequence 755, App
40	47	43.5	206	13	US-08-993-002A-9639	Sequence 9639, Ap
41	47	43.5	78	15	US-09-134-001C-3647	Sequence 3647, Ap
42	46	42.6	940	24	US-60-294-411-11	Sequence 11, Appl
43	46	42.6	40	1	PCT-US98-13684-215	Sequence 215, App
44	45	41.7	40	16	US-09-227-357-215	Sequence 215, App
45	45	41.7				

ALIGNMENTS

RESULT 1

US-09-763-397A-25

; Sequence 25, Application US/09763397A

; GENERAL INFORMATION:

; APPLICANT: The Government of the United States of America, as represented by the

; APPLICANT: Secretary of the Department of Health and Human Services, Centers for

; APPLICANT: Control and Prevention

; APPLICANT: Lal, Altaf A.

; APPLICANT: Hasnain, Seyed E.

; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa

; FILE REFERENCE: 6395-57049

; CURRENT APPLICATION NUMBER: US/09/763,397A

; CURRENT FILING DATE: 2001-02-16

; PRIOR APPLICATION NUMBER: US 60/097,703

; PRIOR FILING DATE: 1998-08-21

; PRIOR APPLICATION NUMBER: PCT / US99/18869

; PRIOR FILING DATE: 1999-08-19

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25

; LENGTH: 22

; TYPE: PRT

; ORGANISM: Honey bee

US-09-763-397A-25

Query Match 100.0%; Score 108; DB 21; Length 22;

Best Local Similarity 100.0%; Pred. No. 4.8e-07; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0;

Qy 1 MKFLVNLVFMVVIYSIYAD 22

Db 1 MKFLNVNVALFVMVYISYIYAD 22

RESULT 2

US-60-292-638-6  
; Sequence 6, Application US/60292638  
; GENERAL INFORMATION:  
; APPLICANT: Morsey, Mohamad A.  
; TITLE OF INVENTION: Non-anaphylactogenic IgE vaccines  
; FILE REFERENCE: PC11011A  
; CURRENT APPLICATION NUMBER: US/60/292,638  
; CURRENT FILING DATE: 2001-05-22  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 129  
; TYPE: PRT  
; ORGANISM: Baculovirus expressed human CH3 domain  
US-60-292-638-6

Query Match 100.0%; Score 108; DB 24; Length 129;  
Best Local Similarity 100.0%; Pred. No. 1.9e-06;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKFLNVNVALFVMVYISYIYAD 22  
Db 1 MKFLNVNVALFVMVYISYIYAD 22

RESULT 3

US-09-763-397A-2  
; Sequence 2, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Recombinant DNA/Protein  
US-09-763-397A-2

Query Match 100.0%; Score 108; DB 21; Length 350;  
Best Local Similarity 100.0%; Pred. No. 4.1e-06;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKFLNVNVALFVMVYISYIYAD 22  
Db 1 MKFLNVNVALFVMVYISYIYAD 22

RESULT 4

US-08-911-312-34  
; Sequence 34, Application US/08911312  
; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William  
; TITLE OF INVENTION: Telomerase Reverse Transcriptase  
; NUMBER OF SEQUENCES: 170  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/911,312  
; FILING DATE: 14-AUG-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/912,951  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/915,503  
; FILING DATE: 14-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Einhorn, Gregory P.  
; REGISTRATION NUMBER: 38,440  
; REFERENCE/DOCKET NUMBER: 015389-002500US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1189 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-911-312-34

Query Match 100.0%; Score 108; DB 13; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKFLNVNVALFVMVYISYIYAD 22  
Db 1 MKFLNVNVALFVMVYISYIYAD 22

RESULT 5

US-08-911-312A-34  
; Sequence 34, Application US/08911312A  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William  
; TITLE OF INVENTION: Telomerase Reverse Transcriptase  
; NUMBER OF SEQUENCES: 171  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/911,312A  
; FILING DATE: 14-AUG-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/912,951  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/915,503  
; FILING DATE: 14-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Einhorn, Gregory P.  
; REGISTRATION NUMBER: 38,440  
; REFERENCE/DOCKET NUMBER: 015389-002500US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1189 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-911-312A-34

Query Match 100.0%; Score 108; DB 13; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLVNLVFMVVIYSIYAD 22  
|||||  
DB 1 MKFLVNLVFMVVIYSIYAD 22

RESULT 6  
US-08-912-951-325  
; Sequence 325, Application US/08912951  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
; TITLE OF INVENTION: THERAPEUTIC METHODS  
; NUMBER OF SEQUENCES: 335  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/912,951  
; FILING DATE: 14-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002600US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 325:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1189 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-912-951-325

Query Match 100.0%; Score 108; DB 13; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKFLVNVALVFEMVYISYIYAD 22  
|||

Dd 1 MKFLVNVALVFEMVYISYIYAD 22  
|||

Db 1 MKFLVNVALVFMVVYISYIYAD 22

## RESULT

```

US-08-974-549-613
; Sequence 613, Application US/08974549
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachimlm
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 726
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 613:
; SEQUENCE CHARACTERISTICS:

```



ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 613:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1189 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1189  
OTHER INFORMATION: /note= "fusion protein composed of  
melittin signal sequence and full length  
hprt protein"  
SEQUENCE DESCRIPTION: SEQ ID NO: 613:  
US-09-432-503-613

Query Match 100.0%; Score 108; DB 18; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKFLVNVALVFMVYIYIYAD 22  
Db 1 MKFLVNVALVFMVYIYIYAD 22

RESULT 9  
US-09-376-330-4  
Sequence 4, Application US/09376330

GENERAL INFORMATION:  
APPLICANT: Tessier, Daniel C.  
APPLICANT: Dignard, Daniel  
APPLICANT: Bergeron, John J.M.  
APPLICANT: Thomas, David Y.  
TITLE OF INVENTION: Method for screening for  
TITLE OF INVENTION: UDP-glucose:glycoprotein glucosyltransferase (UGGT) activity  
FILE REFERENCE: 2139-9"US"  
CURRENT APPLICATION NUMBER: US/09/376,330  
CURRENT FILING DATE: 1999-08-18  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 21  
TYPE: PRT  
ORGANISM: Melittin Signal Peptide sequence  
US-09-376-330-4

Query Match 94.4%; Score 102; DB 17; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKFLVNVALVFMVYIYIYIYA 21  
Db 1 MKFLVNVALVFMVYIYIYIYA 21

RESULT 10  
US-09-766-378-36  
Sequence 36, Application US/09766378  
GENERAL INFORMATION:  
APPLICANT: Rhode, Peter R.  
APPLICANT: Acevedo, Jorge  
APPLICANT: Burkhardt, Martin  
APPLICANT: Jiao, Jin-an

APPLICANT: Wong, Hing C.  
TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND  
METHODS OF USE THEREOF  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: usa  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/766,378  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/960,190  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Corless, Peter F  
REGISTRATION NUMBER: 33,860  
REFERENCE/DOCKET NUMBER: 48002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-766-378-36

Query Match 94.4%; Score 102; DB 21; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKFLVNVALVFMVYIYIYIYA 21  
Db 1 MKFLVNVALVFMVYIYIYIYA 21

RESULT 11  
US-09-766-378A-36  
Sequence 36, Application US/09766378A  
GENERAL INFORMATION:  
APPLICANT: Rhode, Peter R.  
APPLICANT: Acevedo, Jorge  
APPLICANT: Burkhardt, Martin  
APPLICANT: Jiao, Jin-an  
APPLICANT: Wong, Hing C.  
TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND  
METHODS OF USE THEREOF  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: usa  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

```

;
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/766,378A
; FILING DATE: 19-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/960,190
; FILING DATE: 29-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Corless, Peter F
; REGISTRATION NUMBER: 33,860
; REFERENCE/DOCKET NUMBER: 48002-DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-766-378A-36

Query Match          94.4%; Score 102; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLNVNVALFVMVYIYIA 21
   |||||
Db 1 MKFLNVNVALFVMVYIYIA 21
   |||||

RESULT 12
US-09-847-208-26
; Sequence 26, Application US/09847208
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Apis mellifera(Honeybee)Apis cerana(Ind. honeybee)
US-09-847-208-26

Query Match          94.4%; Score 102; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 7.2e-06;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLNVNVALFVMVYIYIA 21
   |||||
Db 1 MKFLNVNVALFVMVYIYIA 21
   |||||

RESULT 13
PCT-US00-02740-5
; Sequence 5, Application PC/TUS0002740
; GENERAL INFORMATION:
; APPLICANT: Johnson et al.
; TITLE OF INVENTION: INHIBITION OF THE MHC CLASS II ANTIGEN PRESENTATION
; TITLE OF INVENTION: PATHWAY AND PRESENTATION TO CD4+ CELLS
US-09-890-806-5

Query Match          94.4%; Score 102; DB 1; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLNVNVALFVMVYIYIA 21
   |||||
Db 1 MKFLNVNVALFVMVYIYIA 21
   |||||

RESULT 14
US-09-890-806-5
; Sequence 5, Application US/09890806
; GENERAL INFORMATION:
; APPLICANT: Johnson et al.
; TITLE OF INVENTION: INHIBITION OF THE MHC CLASS II ANTIGEN PRESENTATION
; TITLE OF INVENTION: PATHWAY AND PRESENTATION TO CD4+ CELLS
; FILE REFERENCE: 0899-59399
; CURRENT APPLICATION NUMBER: US/09/890,806
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: US00/02740
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 60/118,287
; PRIOR FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:chimera
US-09-890-806-5

Query Match          94.4%; Score 102; DB 22; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLNVNVALFVMVYIYIA 21
   |||||
Db 1 MKFLNVNVALFVMVYIYIA 21
   |||||

RESULT 15
US-08-782-481-43
; Sequence 43, Application US/08782481
; GENERAL INFORMATION:
; APPLICANT: Hein, Mich B.
; APPLICANT: Hiatt, Andrew C.
; APPLICANT: Fitchett, John H.
; TITLE OF INVENTION: NOVEL EPITHELIAL TISSUE TARGETING AGENT
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington

```

```

; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/782.481
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 310098.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-782-481-43

```

```

Query Match      86.68; Score 93.5; DB 11; Length 26;
Best Local Similarity 95.5%; Pred. No. 4.3e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

Qy 1 MKFLVNVALFMVYIYIYAD 22
   ||||| ||||| |||||
Db 1 MKFLVNVAL-FMVYIYIYAD 21

```

Search completed: January 29, 2002, 11:06:41  
Job time: 623 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:07:15 ; Search time 29.91 Seconds  
(without alignments)  
26.919 Million cell updates/sec

Title: US-09-763-397A-25

Perfect score: 108  
Sequence: 1 MRFLNVALVFMVYISYIAD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	1189	5	US-09-402-181A-613
2	108	100.0	1189	5	US-09-721-477-613
3	47	43.5	127	5	US-09-815-242-10432
4	45	41.7	39	5	US-09-973-278-211
5	45	41.7	1078	6	US-10-002-854-2
6	43	39.8	63	6	US-10-001-879-128
7	43	39.8	309	5	US-09-886-055-367
8	43	39.8	542	5	US-09-804-291-367
9	43	39.8	509	5	US-09-830-123-2
10	42.5	39.4	717	5	US-09-918-779-12
11	42	38.9	61	5	US-09-959-730-147
12	42	38.9	62	5	US-09-774-639-186
13	42	38.9	217	5	US-09-708-427-19299
14	42	38.9	347	5	US-09-708-427-3624
15	42	38.9	412	5	US-09-708-427-3623
16	42	38.9	456	5	US-09-708-427-17563
17	42	38.9	457	5	US-09-708-427-3622
18	42	38.9	506	5	US-09-708-427-17562
19	42	38.9	569	5	US-09-708-427-17561
20	42	38.9	1297	5	US-09-922-600-17
21	41	38.0	346	5	US-09-886-055-227
22	41	38.0	346	5	US-09-804-291-227
23	40.5	37.5	634	5	US-09-708-427-12942
24	40.5	37.5	678	5	US-09-708-427-12941
25	40.5	37.5	697	5	US-09-708-427-12940
26	40	37.0	92	5	US-09-708-427-62151

27	40	37.0	177	5	US-09-154-707A-328	Sequence 328, App
28	40	37.0	177	5	US-09-966-262-328	Sequence 328, App
29	40	37.0	298	5	US-09-886-055-349	Sequence 349, App
30	40	37.0	298	5	US-09-804-291-349	Sequence 349, App
31	40	37.0	401	5	US-09-965-522-1	Sequence 1, Appli
32	40	37.0	401	5	US-09-991-212-1	Sequence 1, Appli
33	40	37.0	458	5	US-09-897-516-5313	Sequence 5313, Ap
34	40	37.0	493	5	US-09-966-521-84	Sequence 18, Appl
35	40	37.0	502	5	US-09-966-521-84	Sequence 84, Appl
36	40	37.0	540	5	US-09-830-123-15	Sequence 15, Appl
37	40	37.0	553	5	US-09-830-123-17	Sequence 17, Appl
38	40	37.0	555	5	US-09-830-123-19	Sequence 19, Appl
39	40	37.0	621	6	US-10-005-211-2	Sequence 2, Appli
40	39	36.1	62	5	US-09-708-427-73229	Sequence 73229, A
41	39	36.1	175	5	US-09-989-442-108	Sequence 108, App
42	39	36.1	346	5	US-09-886-041-2	Sequence 2, Appli
43	39	36.1	346	5	US-09-995-543-24	Sequence 24, Appl
44	39	36.1	350	5	US-09-968-433-16	Sequence 16, Appl
45	39	36.1	350	5	US-09-876-252-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1  
US-09-402-181A-613  
; Sequence 613, Application US/09402181A  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Lingner, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.  
; Morin, Gregg B.  
; Harley, Calvin B.  
; Andrews, William H.  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 633  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/402,181A  
; FILING DATE: 29-Sep-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/911,312  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: US 08/912,951  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: US 08/915,503  
; FILING DATE: 14-AUG-1997  
; APPLICATION NUMBER: WO PCT/US97/17885  
; FILING DATE: 01-OCT-1997  
; ATTORNEY/AGENT INFORMATION:

NAME: Ausenhuss, Scott L.  
REGISTRATION NUMBER: 42,271  
REFERENCE/DOCKET NUMBER: 015389-00262005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 613:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1189 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1189  
OTHER INFORMATION: /note= "fusion protein composed of  
melittin signal sequence and full length  
hTERT protein"  
SEQUENCE DESCRIPTION: SEQ ID NO: 613:  
US-09-402-181A-613

Query Match 100.0%; Score 108; DB 5; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 5.3e-08;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLVNVALFMVWVYISYIYAD 22  
|||||  
DB 1 MKFLVNVALFMVWVYISYIYAD 22

RESULT 2  
US-09-721-477-613  
Sequence 613, Application US/09721477  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin B.  
Andrews, William H.  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
NUMBER OF SEQUENCES: 727  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/721,477  
FILING DATE: 22-Nov-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/374,549  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
APPLICATION NUMBER: WO PCT/US97/17895  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-00261005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 613:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1189 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1189  
OTHER INFORMATION: /note= "fusion protein composed of  
melittin signal sequence and full length  
hTERT protein"  
SEQUENCE DESCRIPTION: SEQ ID NO: 613:  
US-09-721-477-613

Query Match 100.0%; Score 108; DB 5; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 5.3e-08;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLVNVALFMVWVYISYIYAD 22  
|||||  
DB 1 MKFLVNVALFMVWVYISYIYAD 22

RESULT 3  
US-09-815-242-10432  
Sequence 10432, Application US/09815242  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10432  
LENGTH: 127  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-813-242-10432

Query Match 43.5%; Score 47; DB 5; Length 127;  
Best Local Similarity 40.9%; Pred. No. 3.5;  
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 MKFLNVALFVMVYISYIAD 22  
||:| |||: : :|:|  
Db 17 MKWVVVALLVAIVGNLYRD 38

RESULT 4  
US-09-973-278-211  
Sequence 211, Application US/09973278  
GENERAL INFORMATION:  
APPLICANT: Fischer et al.  
TITLE OF INVENTION: 123 Human Secreted Proteins  
FILE REFERENCE: P2010P2  
CURRENT APPLICATION NUMBER: US/09/973,278  
CURRENT FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: 60/239,899  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 09/227,357  
PRIOR FILING DATE: 1999-01-08  
PRIOR APPLICATION NUMBER: PCT/US98/13684  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/051,926  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/052,793  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,925  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,929  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/052,803  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/052,732  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,931  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,932  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,916  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,930  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,918  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,920  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/052,733  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/052,795  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,919  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,928  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/055,722  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,723  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,948  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,949  
PRIOR FILING DATE: 1997-08-18

PRIOR APPLICATION NUMBER: 60/055,953  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,950  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,947  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,964  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/056,360  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,684  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,984  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,954  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/058,785  
PRIOR FILING DATE: 1997-09-12  
PRIOR APPLICATION NUMBER: 60/058,664  
PRIOR FILING DATE: 1997-09-12  
PRIOR APPLICATION NUMBER: 60/058,660  
PRIOR FILING DATE: 1997-09-12  
PRIOR APPLICATION NUMBER: 60/058,661  
PRIOR FILING DATE: 1997-09-12  
NUMBER OF SEQ ID NOS: 947  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 211  
LENGTH: 39  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-973-278-211

Query Match 41.7%; Score 45; DB 5; Length 39;  
Best Local Similarity 50.0%; Pred. No. 2;  
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 FLNVALFVMVYISYIY 20  
|| | |||: :|:|  
Db 6 FLGNLLVFLILASSFIY 23

RESULT 5  
US-10-002-854-2  
Sequence 2, Application US/10002854  
GENERAL INFORMATION:  
APPLICANT: Scadden, David T.  
APPLICANT: Pozhansky, Mark C.  
APPLICANT: Olszak, Ivona T.  
APPLICANT: Brown, Edward M.  
TITLE OF INVENTION: THE CAR RECEPTOR AS A MEDIATOR OF  
FILE REFERENCE: M0765/7038/ERG/KA  
CURRENT APPLICATION NUMBER: US/10/002,854  
CURRENT FILING DATE: 2001-11-01  
PRIOR APPLICATION NUMBER: PCT/US00/15440  
PRIOR FILING DATE: 2000-06-02  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 1078  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)...(1078)  
OTHER INFORMATION: Ca-sensing Receptor  
US-10-002-854-2

Query Match 41.7%; Score 45; DB 6; Length 1078;  
Best Local Similarity 45.0%; Pred. No. 60;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```
QY 2 KFLVNVLFVWVYISYIYA 21
||: |::|::|::|::|
Db 805 KFITFSMLFIWFVWISFIPA 824

RESULT 6
US-10-001-879-128
; Sequence 128, Application US/10001879
; GENERAL INFORMATION:
; APPLICANT: Saiceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Heirve
; APPLICANT: Cafferkey, Robert
; APPLICANT: Ali, Shujath
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0281
; CURRENT APPLICATION NUMBER: US/10/001,879
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,188
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 128
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-879-128

Query Match 39.8%; Score 43; DB 6; Length 63;
Best Local Similarity 44.4%; Pred. No. 6.5;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 FLVNVLFVWVYISYIY 20
||| |::|::|::|::|
Db 4 FLVLSLIYVMEYIIFFF 21

RESULT 7
US-09-886-055-367
; Sequence 367, Application US/09886055
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 367
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-367

Query Match 39.8%; Score 43; DB 5; Length 309;
Best Local Similarity 40.0%; Pred. No. 33;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 6 NVALVFMVYISYIY 20
|: |::|::|::|::|
Db 204 NIFFVLLVIFISYLF 218

RESULT 8
US-09-886-055-367
; Sequence 367, Application US/09886055
; GENERAL INFORMATION:
; APPLICANT: ZOZULYA, SERGEY
; APPLICANT: STRYER, LUBERT
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 367
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-367

Query Match 39.8%; Score 43; DB 5; Length 309;
Best Local Similarity 40.0%; Pred. No. 33;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 6 NVALVFMVYISYIY 20
|: |::|::|::|::|
Db 204 NIFFVLLVIFISYLF 218

RESULT 9
US-09-830-123-2
; Sequence 2, Application US/09830123
; GENERAL INFORMATION:
; APPLICANT: Iida, Shigeru
; APPLICANT: Tanaka, Sachiko
; APPLICANT: Inagaki, Yoshihide
; TITLE OF INVENTION: Genes Encoding Proteins Regulating the pH of Vacuoles
; FILE REFERENCE: 001560-397
; CURRENT APPLICATION NUMBER: US/09/830,123
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/Jp00/05722
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: JP 11/236800
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Ipomea nil
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(542)
; OTHER INFORMATION: Amino acid sequence of protein regulating the pH of vacuoles
US-09-830-123-2

Query Match 39.8%; Score 43; DB 5; Length 542;
```







Qy 7 VALVEMVYISYIYAD 22  
| | | : : | : | : | :  
Db 226 VALMMLMAYLSYMLAE 241

Search completed: January 29, 2002, 11:07:16  
Job time: 534 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:06:40 ; Search time 353.23 Seconds  
(without alignments)  
11.791 Million cell updates/sec

Title: us-09-763-397A-24

Perfect score: 74

Sequence: 1 QYIKANSKFIGITEL 15

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCRUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	74	100.0	15	1	PCT-US00-25239-7	Sequence 7, Appli
2	74	100.0	15	1	PCT-US99-06325-2	Sequence 2, Appli
3	74	100.0	15	5	US-08-161-889-69	Sequence 69, Appl
4	74	100.0	15	5	US-08-161-889A-69	Sequence 69, Appl
5	74	100.0	15	8	US-08-406-916B-2	Sequence 2, Appli
6	74	100.0	15	11	US-08-716-249-4	Sequence 4, Appli
7	74	100.0	15	14	US-09-049-847-1	Sequence 1, Appli
8	74	100.0	15	14	US-09-049-847A-1	Sequence 1, Appli
9	74	100.0	15	15	US-09-102-716-6	Sequence 6, Appli

10	74	100.0	15	15	US-09-148-711-7	Sequence 7, Appli
11	74	100.0	15	15	US-09-148-711A-7	Sequence 7, Appli
12	74	100.0	15	15	US-09-165-878-71	Sequence 71, Appl
13	74	100.0	15	17	US-09-362-731-13	Sequence 13, Appl
14	74	100.0	15	17	US-09-364-088-6	Sequence 6, Appli
15	74	100.0	15	17	US-09-396-937-34	Sequence 34, Appl
16	74	100.0	15	18	US-09-405-986-1	Sequence 1, Appli
17	74	100.0	15	18	US-09-413-186-12	Sequence 12, Appl
18	74	100.0	15	19	US-09-523-279-6	Sequence 6, Appli
19	74	100.0	15	19	US-09-556-818-23	Sequence 23, Appl
20	74	100.0	15	19	US-09-580-015-49	Sequence 49, Appl
21	74	100.0	15	19	US-09-580-018-49	Sequence 49, Appl
22	74	100.0	15	19	US-09-580-019-49	Sequence 49, Appl
23	74	100.0	15	21	US-09-743-731-2	Sequence 2, Appli
24	74	100.0	15	21	US-09-763-397A-24	Sequence 24, Appl
25	74	100.0	15	21	US-09-785-215-4	Sequence 4, Appli
26	74	100.0	15	22	US-09-862-849-2	Sequence 2, Appli
27	74	100.0	16	1	PCT-US99-03055-55	Sequence 55, Appl
28	74	100.0	16	9	US-08-577-106-1	Sequence 1, Appli
29	74	100.0	16	14	US-09-091-466-4	Sequence 4, Appli
30	74	100.0	16	15	US-09-165-878-18	Sequence 18, Appl
31	74	100.0	16	22	US-09-848-834A-2	Sequence 2, Appli
32	74	100.0	17	1	PCT-US94-04832A-4	Sequence 4, Appli
33	74	100.0	17	1	PCT-US99-13923-32	Sequence 32, Appl
34	74	100.0	17	1	PCT-US99-13959-63	Sequence 63, Appl
35	74	100.0	17	4	US-08-057-166-4	Sequence 4, Appli
36	74	100.0	17	4	US-08-060-798A-7	Sequence 7, Appli
37	74	100.0	17	6	US-08-218-461-7	Sequence 7, Appli
38	74	100.0	17	6	US-08-229-275-4	Sequence 4, Appli
39	74	100.0	17	6	US-08-245-507-12	Sequence 12, Appl
40	74	100.0	17	7	US-08-328-912B-7	Sequence 7, Appli
41	74	100.0	17	8	US-08-472-701-23	Sequence 23, Appl
42	74	100.0	17	8	US-08-472-704-23	Sequence 23, Appl
43	74	100.0	17	8	US-08-488-320A-4	Sequence 4, Appli
44	74	100.0	17	11	US-08-718-490A-7	Sequence 7, Appli
45	74	100.0	17	13	US-08-926-296-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1  
PCT-US00-25239-7  
; Sequence 7, Application PC/TUS0025239  
; GENERAL INFORMATION:  
; APPLICANT: Schenk, Dale B  
; APPLICANT: Elan Pharmaceuticals  
; TITLE OF INVENTION: Prevention and Treatment of Amyloidogenic Disease  
; FILE REFERENCE: 15270J-005900PC  
; CURRENT APPLICATION NUMBER: PCT/US00/25239  
; CURRENT FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: US 60/137,010  
; PRIOR FILING DATE: 1999-06-01  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:tetanus toxoid  
; OTHER INFORMATION: TT 830-844 universal T-cell epitope  
PCT-US00-25239-7

Query Match 100.0%; Score 74; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 QYIKANSKFIGITEL 15  
Db 1 QYIKANSKFIGITEL 15



```
;
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLARD, SIDNEY W.
; STREET: 7632 SLATE RIDGE BOULEVARD
; CITY: REYNOLDSBURG
; STATE: OHIO
; COUNTRY: USA
; ZIP: 43068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PS/2 COMPATIBLE 486 DX 33
; OPERATING SYSTEM: MS DOS 6.2
; SOFTWARE: WORDPERFECT FOR MS DOS 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/406.916B
; FILING DATE: 27 MAR 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08370
; FILING DATE: 30 SEP 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KREMBLAS, FRANCIS T., JR.
; REGISTRATION NUMBER: 22,773
; REFERENCE/DOCKET NUMBER: URF 2 056 3 3 3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614) 575 2100
; TELEFAX: (614) 575 2149
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: peptide
; HYPOTHEICAL: No
; FRAGMENT TYPE: 830-844 fragment of tetanus toxoid
; FEATURE:
; OTHER INFORMATION: T cell epitope
; US-08-406-916B-2

Query Match 100.0%; Score 74; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 6
US-08-716-249-4
; Sequence 4, Application US/08716249
; GENERAL INFORMATION:
; APPLICANT: Guichard, Gilles
; APPLICANT: Muller, Sylviane
; APPLICANT: Briand, Jean-Paul
; APPLICANT: Regenmortel, Marc
; TITLE OF INVENTION: Retropeptides, Antibodies Thereto, and
; TITLE OF INVENTION: Uses Thereof for Vaccination and In Vitro Diagnosis
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spencer & Frank
; STREET: 1100 New York Avenue, Suite 300E
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716.249
```

```
;
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00292
; FILING DATE: 13-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Calvetti, Frederick F.
; REGISTRATION NUMBER: 28,557
; REFERENCE/DOCKET NUMBER: GROFO 7001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)414-4000
; TELEFAX: (202)414-4040
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-716-249-4

Query Match 100.0%; Score 74; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 7
US-09-049-847-1
; Sequence 1, Application US/09049847
; GENERAL INFORMATION:
; APPLICANT: Bay, Sylvie
; APPLICANT: Cantacuzene, Daniele
; APPLICANT: Leclerc, Claude
; APPLICANT: Lo-Man, Richard
; TITLE OF INVENTION: Multiple antigen alycopeptide carbohydrate, vaccine
; TITLE OF INVENTION: comprising the same and use thereof
; FILE REFERENCE: 102.166A
; CURRENT APPLICATION NUMBER: US/09/049,847
; CURRENT FILING DATE: 1998-03-27
; EARLIER APPLICATION NUMBER: 60/041,726
; EARLIER FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
; US-09-049-847-1
```

```
Query Match 100.0%; Score 74; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 8
US-09-049-847A-1
; Sequence 1, Application US/09049847A
; GENERAL INFORMATION:
; APPLICANT: Bay, Sylvie
; APPLICANT: Cantacuzene, Daniele
; APPLICANT: Leclerc, Claude
; APPLICANT: Lo-Man, Richard
```

; TITLE OF INVENTION: Mutiple antigen glycopeptide carbohydrate, vaccine

; FILE REFERENCE: 102.166A

; CURRENT APPLICATION NUMBER: US/09/049,847A

; CURRENT FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/041,726

; PRIOR FILING DATE: 1997-03-27

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 15

; TYPE: PRT

; ORGANISM: Clostridium tetani

US-09-049-847A-1

Query Match

Best Local Similarity 100.0%; Score 74; DB 14; Length 15;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15

Db 1 QYIKANSKFIGITEL 15

RESULT 9

US-09-102-716-6

; Sequence 6, Application US/09102716

; GENERAL INFORMATION:

; APPLICANT: Yashwant M. Deo

; Joel Goldstein

; Robert Graziano

; Chezian Somasundaram

; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED

; OF ANTI-FC RECEPTOR ANTIBODIES

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street, Suite 510

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109-1875

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/102,716

; FILING DATE: 22-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/661,052

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Arnold, Beth E.

; REGISTRATION NUMBER: 35,430

; REFERENCE/DOCKET NUMBER: MXI-043CP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)227-7400

; TELEFAX: (617)227-5941

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 15 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-102-716-6

Query Match

Best Local Similarity 100.0%; Score 74; DB 15; Length 15;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 1.2e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15

Db 1 QYIKANSKFIGITEL 15

RESULT 10

US-09-148-711-7

; Sequence 7, Application US/09148711

; GENERAL INFORMATION:

; APPLICANT: Bakaletz, Lauren O.

; APPLICANT: Kaumaya, Parvin T.

; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Calfee, Halter and Griswold

; STREET: 800 Superior Avenue

; CITY: Cleveland

; STATE: Ohio

; COUNTRY: U.S.A.

; ZIP: 44114-2688

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/148,711

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/460,502

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Goltick, Mary E.

; REGISTRATION NUMBER: 34,829

; REFERENCE/DOCKET NUMBER: 22727/00120

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (216) 622-8458

; TELEFAX: (216) 241-0816

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 15 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-148-711-7

Query Match 100.0%; Score 74; DB 15; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.2e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15

Db 1 QYIKANSKFIGITEL 15

RESULT 11

US-09-148-711A-7

; Sequence 7, Application US/09148711A

; GENERAL INFORMATION:

; APPLICANT: The Ohio State University

; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides

; FILE REFERENCE: 18525-04010

; CURRENT APPLICATION NUMBER: US/09/148,711A

; CURRENT FILING DATE: 1998-09-04

; PRIOR APPLICATION NUMBER: 08/460,502

; PRIOR FILING DATE: 1995-06-02

; NUMBER OF SEQ ID NOS: 12



; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: clostridium tetani  
US-09-148-711A-7

Query Match 100.0%; Score 74; DB 15; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15  
|||||

Db 1 QYIKANSKFIGITEL 15  
|||||

## RESULT 12

US-09-165-878-71  
; Sequence 71, Application US/09165878B  
; GENERAL INFORMATION:  
; APPLICANT: Heegaard, Peter Mikael Helweg  
; APPLICANT: Jakobsen, Palle Hoy  
; TITLE OF INVENTION: Non-Dendritic Backbone Peptide Carrier  
; FILE REFERENCE: 2316.1009-000  
; CURRENT APPLICATION NUMBER: US/09/165,878B  
; CURRENT FILING DATE: 1998-10-02  
; EARLIER APPLICATION NUMBER: PCT/DK97/00146  
; EARLIER FILING DATE: 1997-04-03  
; EARLIER APPLICATION NUMBER: DK 0398/96  
; EARLIER FILING DATE: 1996-04-03  
; NUMBER OF SEQ ID NOS: 113  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 71  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Unknown  
US-09-165-878-71

Query Match 100.0%; Score 74; DB 15; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15  
|||||

Db 1 QYIKANSKFIGITEL 15  
|||||

## RESULT 13

US-09-362-731-13  
; Sequence 13, Application US/09362731  
; GENERAL INFORMATION:  
; APPLICANT: U.C.B. S.A.  
; TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE  
; TREATMENT OF ALLERGY  
; FILE REFERENCE: P.UB.09/WO  
; CURRENT APPLICATION NUMBER: US/09/362,731  
; CURRENT FILING DATE: 1999-07-29  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-362-731-13

Query Match 100.0%; Score 74; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15  
|||||

Db 1 QYIKANSKFIGITEL 15  
|||||

## RESULT 14

US-09-364-088-6  
; Sequence 6, Application US/09364088  
; GENERAL INFORMATION:  
; APPLICANT: Yashwant M. Deo, et al.  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED  
; OF ANTI-FC RECEPTOR ANTIBODIES  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street, 24th Floor  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/364,088  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/188,082  
; FILING DATE: 07-JUNE-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/484,172  
; FILING DATE: 07-JUNE-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remillard, Jane E.  
; REGISTRATION NUMBER: 38,872  
; REFERENCE/DOCKET NUMBER: MX1-043CP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-7414  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: Internal  
US-09-364-088-6

Query Match 100.0%; Score 74; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15  
|||||

Db 1 QYIKANSKFIGITEL 15  
|||||

## RESULT 15

US-09-396-937-34  
; Sequence 34, Application US/09396937  
; GENERAL INFORMATION:  
; APPLICANT: M&E Biotech A/S  
; APPLICANT: HAKKIER, Torben  
; APPLICANT: HAANKING, Jesper  
; TITLE OF INVENTION: Method for Down-Regulating Osteoprotegerin Ligand  
; ACTIVITY

; FILE REFERENCE: 22021 PC 1  
; CURRENT APPLICATION NUMBER: US/09/396,937  
; CURRENT FILING DATE: 1999-09-15  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-09-396-937-34

Query Match 100.0%; Score 74; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred No. 1.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15  
| | | | | | | | | | | | | | | |  
Db 1 QYIKANSKFIGITEL 15

Search completed: January 29, 2002, 11:06:40  
Job time: 622 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:07:15 ; Search time 29.91 Seconds  
(without alignments)  
18.354 Million cell updates/sec

Title: US-09-763-397A-24

Perfect score: 74

Sequence: 1 QYKANSKFIGITEL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	100.0	15	5	US-09-787-126-34
2	74	100.0	15	5	US-09-677-374-19
3	74	100.0	15	5	US-09-806-703A-12
4	74	100.0	15	5	US-09-980-916-23
5	74	100.0	25	5	US-09-806-703A-15
6	74	100.0	25	5	US-09-806-703A-16
7	74	100.0	25	5	US-09-806-703A-17
8	74	100.0	116	5	US-09-980-916-14
9	74	100.0	116	5	US-09-980-916-52
10	74	100.0	118	5	US-09-980-916-3
11	74	100.0	118	5	US-09-980-916-36
12	74	100.0	122	5	US-09-980-916-15
13	74	100.0	122	5	US-09-980-916-16
14	74	100.0	122	5	US-09-980-916-54
15	74	100.0	124	5	US-09-980-916-4
16	74	100.0	124	5	US-09-980-916-5
17	74	100.0	124	5	US-09-980-916-13
18	74	100.0	124	5	US-09-980-916-17
19	74	100.0	124	5	US-09-980-916-34
20	74	100.0	124	5	US-09-980-916-38
21	74	100.0	124	5	US-09-980-916-46
22	74	100.0	126	5	US-09-980-916-2
23	74	100.0	126	5	US-09-980-916-6
24	74	100.0	126	5	US-09-980-916-28
25	74	100.0	139	5	US-09-980-916-22
26	74	100.0	141	5	US-09-980-916-11

27	74	100.0	145	5	US-09-980-916-60	Sequence 60, Appl
28	74	100.0	147	5	US-09-980-916-44	Sequence 44, Appl
29	74	100.0	173	5	US-09-787-126-18	Sequence 18, Appl
30	74	100.0	182	5	US-09-787-126-16	Sequence 16, Appl
31	74	100.0	199	5	US-09-806-703A-39	Sequence 39, Appl
32	74	100.0	208	5	US-09-806-703A-37	Sequence 37, Appl
33	68	91.9	19	5	US-09-909-460-41	Sequence 41, Appl
34	66	89.2	14	5	US-09-834-784-510	Sequence 510, Appl
35	65	87.8	13	5	US-09-909-460-42	Sequence 42, Appl
36	46	62.2	877	5	US-09-815-242-4901	Sequence 4901, Ap
37	46	62.2	880	5	US-09-815-242-10491	Sequence 10491, A
38	42	56.8	455	5	US-09-897-516-6808	Sequence 6808, Ap
39	40	54.1	644	5	US-09-946-290-2	Sequence 2, Appl
40	38	51.4	285	5	US-09-897-516-5943	Sequence 5943, Ap
41	37	50.0	317	5	US-09-897-516-5603	Sequence 5603, Ap
42	36	48.6	376	5	US-09-978-756-2	Sequence 2, Appl
43	36	48.6	594	5	US-09-117-4158-2	Sequence 2, Appl
44	36	48.6	613	5	US-09-117-4158-22	Sequence 22, Appl
45	36	48.6	631	5	US-09-117-4158-18	Sequence 18, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-787-126-34  
; Sequence 34, Application US/09787126  
; GENERAL INFORMATION:  
; APPLICANT: M&E Biotech A/S  
; APPLICANT: HALKIER, Torben  
; APPLICANT: HAANING, Jesper  
; TITLE OF INVENTION: Method for Down-Regulating Osteoprotegerin Ligand  
; TITLE OF INVENTION: Activity  
; FILE REFERENCE: 3631-0108P  
; CURRENT APPLICATION NUMBER: US/09/787,126  
; CURRENT FILING DATE: 2001-10-09  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-09-787-126-34

Query Match 100.0%; Score 74; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYKANSKFIGITEL 15

Db 1 QYKANSKFIGITEL 15

##### RESULT 2

US-09-677-374-19  
; Sequence 19, Application US/09677374  
; GENERAL INFORMATION:  
; APPLICANT: Burian, Jan  
; APPLICANT: Kuzyk, Michael  
; APPLICANT: Thornton, Julian  
; APPLICANT: Kay, William  
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY AGAINST  
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY  
; FILE REFERENCE: IDC01/60485/US  
; CURRENT APPLICATION NUMBER: US/09/677,374  
; CURRENT FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/154,437  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: NO 20004637  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: IE 2000/0752  
; PRIOR FILING DATE: 2000-09-18

RESULT 6  
US-09-806-703A-16  
; Sequence 16, Application US/09806703A  
; GENERAL INFORMATION:  
; APPLICANT: Steinaa, Lucilla  
; APPLICANT: Mouritsen, Soren  
; APPLICANT: Gautam, Anand  
; APPLICANT: Dalum, Iben  
; APPLICANT: Haaning, Jesper  
; APPLICANT: Leach, Dana  
; APPLICANT: Nielsen, Klaus

RESULT 4  
US-09-980-916-23  
; Sequence 23, Application US/09980916  
; GENERAL INFORMATION:  
; APPLICANT: M&E Biotech A/S  
; APPLICANT: Klysner, Steen  
; TITLE OF INVENTION: Method for Down-Regulating IL5 Activity  
; FILE REFERENCE: 3631-0112P  
; CURRENT APPLICATION NUMBER: US/09/980, 916  
; CURRENT FILING DATE: 2001-10-23

APPLICANT: Karlsson, Gunilla  
APPLICANT: Rasmussen, Peter  
TITLE OF INVENTION: Novel Methods for Therapeutic Vaccination  
FILE REFERENCE: 3631-0109P  
CURRENT APPLICATION NUMBER: US/09/806,703A  
CURRENT FILING DATE: 2001-04-04  
PRIOR APPLICATION NUMBER: PCT/DK99/00525  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: DK 1998 01261  
PRIOR FILING DATE: 1998-10-05  
PRIOR APPLICATION NUMBER: US 60/105,011  
PRIOR FILING DATE: 1998-10-20  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn Ver. 3.0  
SEQ ID NO 16  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Fusion of  
OTHER INFORMATION: tetanus toxoid epitope and PSM  
US-09-806-703A-16

Query Match 100.0%; Score 74; DB 5; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15  
DB 6 QYIKANSKFIGITEL 20

RESULT 7  
US-09-806-703A-17  
Sequence 17, Application US/09806703A  
GENERAL INFORMATION:  
APPLICANT: Steinnaa, Lucilla  
APPLICANT: Mouritsen, Soren  
APPLICANT: Gautam, Anand  
APPLICANT: Dalum, Iben  
APPLICANT: Haaning, Jesper  
APPLICANT: Leach, Dana  
APPLICANT: Nielsen, Klaus  
APPLICANT: Karlsson, Gunilla  
APPLICANT: Rasmussen, Peter  
TITLE OF INVENTION: Novel Methods for Therapeutic Vaccination  
FILE REFERENCE: 3631-0109P  
CURRENT APPLICATION NUMBER: US/09/806,703A  
CURRENT FILING DATE: 2001-04-04  
PRIOR APPLICATION NUMBER: PCT/DK99/00525  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: DK 1998 01261  
PRIOR FILING DATE: 1998-10-05  
PRIOR APPLICATION NUMBER: US 60/105,011  
PRIOR FILING DATE: 1998-10-20  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn Ver. 3.0  
SEQ ID NO 17  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Fusion of  
OTHER INFORMATION: tetanus toxoid epitope and PSM  
US-09-806-703A-17

Query Match 100.0%; Score 74; DB 5; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15

DB 6 QYIKANSKFIGITEL 20

RESULT 8  
US-09-980-916-14  
Sequence 14, Application US/09980916  
GENERAL INFORMATION:  
APPLICANT: M&E Biotech A/S  
APPLICANT: Klysner, Steen  
TITLE OF INVENTION: Method For Down-Regulating IL5 Activity  
FILE REFERENCE: 3631-0112P  
CURRENT APPLICATION NUMBER: US/09/980,916  
CURRENT FILING DATE: 2001-10-23  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 14  
LENGTH: 116  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Murine IL5  
OTHER INFORMATION: modified by substitution with tetanus toxoid P2  
OTHER INFORMATION: epitope  
NAME/KEY: MUTAGEN  
LOCATION: (30)..(44)  
OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO: 23)  
NAME/KEY: SIMILAR  
LOCATION: (1)..(29)  
OTHER INFORMATION: Identical to residues 1-29 in SEQ ID NO: 12  
NAME/KEY: SIMILAR  
LOCATION: (45)..(116)  
OTHER INFORMATION: Identical to residues 42-113 in SEQ ID NO: 12  
US-09-980-916-14

Query Match 100.0%; Score 74; DB 5; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15  
DB 30 QYIKANSKFIGITEL 44

RESULT 9  
US-09-980-916-52  
Sequence 52, Application US/09980916  
GENERAL INFORMATION:  
APPLICANT: M&E Biotech A/S  
APPLICANT: Klysner, Steen  
TITLE OF INVENTION: Method For Down-Regulating IL5 Activity  
FILE REFERENCE: 3631-0112P  
CURRENT APPLICATION NUMBER: US/09/980,916  
CURRENT FILING DATE: 2001-10-23  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 52  
LENGTH: 116  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Murine IL-5  
OTHER INFORMATION: modified by substitution with tetanus toxoid  
OTHER INFORMATION: epitope  
US-09-980-916-52

Query Match 100.0%; Score 74; DB 5; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15

Db 30 QYIKANSKFIGITEL 44  
|||||

## RESULT 10

US-09-980-916-3  
; Sequence 3, Application US/09980916  
; GENERAL INFORMATION:  
; APPLICANT: M&E Biotech A/S  
; APPLICANT: Klysner, Steen  
; TITLE OF INVENTION: Method For Down-Regulating IL5 Activity  
; FILE REFERENCE: 3631-0112P  
; CURRENT APPLICATION NUMBER: US/09/980,916  
; CURRENT FILING DATE: 2001-10-23  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 118

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Human IL5  
; OTHER INFORMATION: modified by substitution with tetanus toxoid P2

; OTHER INFORMATION: epitope

NAME/KEY: MUTAGEN

LOCATION: (32)..(46)

; OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO: 23)

NAME/KEY: SIMILAR

LOCATION: (1)..(31)

; OTHER INFORMATION: Identical to residues 1-31 in SEQ ID NO: 1

NAME/KEY: SIMILAR

LOCATION: (47)..(118)

; OTHER INFORMATION: Identical to residues 44-115 in SEQ ID NO: 1

US-09-980-916-3

Query Match

100.0%; Score 74; DB 5; Length 118;

Best Local Similarity 100.0%; Pred. No. 1.2e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15  
|||||Db 32 QYIKANSKFIGITEL 46  
|||||

## RESULT 11

US-09-980-916-36  
; Sequence 36, Application US/09980916  
; GENERAL INFORMATION:  
; APPLICANT: M&E Biotech A/S  
; APPLICANT: Klysner, Steen  
; TITLE OF INVENTION: Method For Down-Regulating IL5 Activity  
; FILE REFERENCE: 3631-0112P  
; CURRENT APPLICATION NUMBER: US/09/980,916  
; CURRENT FILING DATE: 2001-10-23  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 36  
; LENGTH: 118

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Human IL-5  
; OTHER INFORMATION: modified by substitution with tetanus toxoid

; OTHER INFORMATION: epitope

US-09-980-916-36

Query Match

100.0%; Score 74; DB 5; Length 118;

Best Local Similarity 100.0%; Pred. No. 1.2e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15  
|||||

Db 32 QYIKANSKFIGITEL 46  
|||||

## RESULT 12

US-09-980-916-15  
; Sequence 15, Application US/09980916  
; GENERAL INFORMATION:  
; APPLICANT: M&E Biotech A/S  
; APPLICANT: Klysner, Steen  
; TITLE OF INVENTION: Method For Down-Regulating IL5 Activity  
; FILE REFERENCE: 3631-0112P  
; CURRENT APPLICATION NUMBER: US/09/980,916  
; CURRENT FILING DATE: 2001-10-23  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 122

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Murine IL5  
; OTHER INFORMATION: modified by substitution with tetanus toxoid P2

; OTHER INFORMATION: epitope

NAME/KEY: MUTAGEN

LOCATION: (37)..(71)

; OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO: 23)

NAME/KEY: SIMILAR

LOCATION: (1)..(56)

; OTHER INFORMATION: Identical to residues 1-56 in SEQ ID NO: 12

NAME/KEY: SIMILAR

LOCATION: (72)..(122)

; OTHER INFORMATION: Identical to residues 63-113 in SEQ ID NO: 12

US-09-980-916-15

Query Match

100.0%; Score 74; DB 5; Length 122;

Best Local Similarity 100.0%; Pred. No. 1.3e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15  
|||||Db 57 QYIKANSKFIGITEL 71  
|||||

## RESULT 13

US-09-980-916-16  
; Sequence 16, Application US/09980916  
; GENERAL INFORMATION:  
; APPLICANT: M&E Biotech A/S  
; APPLICANT: Klysner, Steen  
; TITLE OF INVENTION: Method For Down-Regulating IL5 Activity  
; FILE REFERENCE: 3631-0112P  
; CURRENT APPLICATION NUMBER: US/09/980,916  
; CURRENT FILING DATE: 2001-10-23  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 122

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Murine IL5  
; OTHER INFORMATION: modified by substitution with tetanus toxoid P2

; OTHER INFORMATION: epitope

NAME/KEY: MUTAGEN

LOCATION: (84)..(98)

; OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO: 23)

NAME/KEY: SIMILAR

LOCATION: (1)..(83)

; OTHER INFORMATION: Identical to residues 1-83 in SEQ ID NO: 12

NAME/KEY: SIMILAR

LOCATION: (99)..(122)

; OTHER INFORMATION: Identical to residues 90-113 in SEQ ID NO: 12  
US-09-980-916-16

Query Match 100.0%; Score 74; DB 5; Length 122;  
Best Local Similarity 100.0%; Pred. No. 1.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15  
| | | | | | | | | | | | | | |  
Db 84 QYIKANSKFIGITEL 98

## RESULT 14

US-09-980-916-54  
; Sequence 54, Application US/09980916  
; GENERAL INFORMATION:  
; APPLICANT: M&E Biotech A/S  
; APPLICANT: Klysner, Steen  
; TITLE OF INVENTION: Method For Down-Regulating IL5 Activity  
; FILE REFERENCE: 3631-0112P  
; CURRENT APPLICATION NUMBER: US/09/980,916  
; CURRENT FILING DATE: 2001-10-23  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 54  
; LENGTH: 122  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Murine IL-5  
; OTHER INFORMATION: modified by substitution with tetanus toxoid  
; OTHER INFORMATION: epitope  
US-09-980-916-54

Query Match 100.0%; Score 74; DB 5; Length 122;  
Best Local Similarity 100.0%; Pred. No. 1.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15  
| | | | | | | | | | | | | | |  
Db 84 QYIKANSKFIGITEL 98

## RESULT 15

US-09-980-916-4  
; Sequence 4, Application US/09980916  
; GENERAL INFORMATION:  
; APPLICANT: M&E Biotech A/S  
; APPLICANT: Klysner, Steen  
; TITLE OF INVENTION: Method For Down-Regulating IL5 Activity  
; FILE REFERENCE: 3631-0112P  
; CURRENT APPLICATION NUMBER: US/09/980,916  
; CURRENT FILING DATE: 2001-10-23  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 124  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Human IL5  
; OTHER INFORMATION: modified by substitution with tetanus toxoid P2  
; OTHER INFORMATION: epitope  
; NAME/KEY: MUTAGEN  
; LOCATION: (59)..(73)  
; OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO:23)  
; NAME/KEY: SIMILAR  
; LOCATION: (1)..(58)  
; OTHER INFORMATION: Identical to residues 1-58 in SEQ ID NO: 1  
; NAME/KEY: SIMILAR  
; LOCATION: (74)..(124)

; OTHER INFORMATION: Identical to residues 65-115 in SEQ ID NO: 1  
US-09-980-916-4

Query Match 100.0%; Score 74; DB 5; Length 124;  
Best Local Similarity 100.0%; Pred. No. 1.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15  
| | | | | | | | | | | | | | |  
Db 59 QYIKANSKFIGITEL 73

Search completed: January 29, 2002, 11:07:15  
Job time: 533 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 11:06:39 ; Search time 353.23 Seconds  
(without alignments)  
13.363 Million cell updates/sec

Title: US-09-763-397A-23

Perfect score: 84

Sequence: 1 LATRLMKFKAEIRDF 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending\_Patents\_AA\_Main:\*
- 1: /cgn2\_6/ptodata/2/paa/PCUTUS\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
  - 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
  - 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
  - 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
  - 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
  - 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
  - 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
  - 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*
  - 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
  - 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
  - 24: /cgn2\_6/ptodata/2/paa/US060\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	84	100.0	17	US-09-763-397A-23
2	84	100.0	21	Sequence 23, Appl
3	45	53.6	743	US-09-763-397A-2
4	44	52.4	169	US-09-248-796-17817
5	44	52.4	174	US-09-583-110-4656
6	44	52.4	225	US-09-107-433-3442
7	43	51.2	62	US-60-269-308-4746
8	43	51.2	86	US-09-417-507-25566
9	43	51.2	86	1 PCT-US00-07726-121
				Sequence 121, Appl
				Sequence 122, Appl

10	51.2	98	1	PCT-US00-05988-1711	Sequence 1711, Ap
11	51.2	98	23	US-09-925-300-1711	Sequence 1711, Ap
12	51.2	155	24	US-60-146-394-1642	Sequence 1642, Ap
13	51.2	301	21	US-09-760-475-1793	Sequence 1793, Ap
14	51.2	316	21	US-09-760-475-3012	Sequence 3012, Ap
15	51.2	509	24	US-60-245-201-221	Sequence 221, App
16	51.2	649	24	US-60-173-464-27800	Sequence 27800, A
17	51.2	649	24	US-60-191-637-38112	Sequence 36112, A
18	51.2	649	24	US-60-191-681-28258	Sequence 28258, A
19	51.2	83	24	US-60-212-664-426	Sequence 426, App
20	51.2	83	24	US-60-216-770-99	Sequence 99, Appl
21	51.2	186	17	US-09-391-631-3853	Sequence 3853, Ap
22	51.2	190	17	US-09-391-631-3852	Sequence 3852, Ap
23	51.2	203	17	US-09-391-631-3851	Sequence 3851, Ap
24	51.2	203	18	US-09-428-944-882	Sequence 882, App
25	48.8	2325	10	US-08-697-826-10	Sequence 10, Appl
26	48.8	2325	10	US-08-697-826A-10	Sequence 10, Appl
27	48.8	2325	16	US-09-237-415-6	Sequence 6, Appl
28	48.8	2325	21	US-09-720-379-6	Sequence 6, Appl
29	47.6	62	20	US-09-614-912-108	Sequence 108, App
30	47.6	62	24	US-60-172-959-14	Sequence 14, Appl
31	47.6	95	1	PCT-US01-08631-36938	Sequence 36938, A
32	47.6	186	1	PCT-US01-08631-36939	Sequence 36939, A
33	47.6	344	24	US-60-324-109-17910	Sequence 17910, A
34	47.6	346	19	US-09-573-655A-1564	Sequence 1564, Ap
35	47.6	534	1	PCT-US99-14689-2	Sequence 2, Appl
36	47.6	535	24	US-60-324-109-17012	Sequence 17012, A
37	47.6	668	16	US-09-252-991A-22341	Sequence 22341, A
38	47.6	775	5	US-08-158-941-15	Sequence 15, Appl
39	47.6	775	5	US-08-158-941-16	Sequence 16, Appl
40	46.4	156	16	US-09-270-767-57787	Sequence 57787, A
41	46.4	156	16	US-09-270-849B-182968	Sequence 182968, A
42	46.4	247	15	US-09-147-729-10	Sequence 10, Appl
43	46.4	247	16	US-09-284-426-16	Sequence 16, Appl
44	46.4	247	20	US-09-632-426-11	Sequence 11, Appl
45	46.4	247	20	US-09-632-570-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-09-763-397A-23

; Sequence 23, Application US/09763397A

; GENERAL INFORMATION:

; APPLICANT: The Government of the United States of America, as represented by the

; APPLICANT: Secretary of the Department of Health and Human Services, Centers for

; APPLICANT: Control and Prevention

; APPLICANT: Lal, Altaf A.

; APPLICANT: Ping Shi, Ya

; APPLICANT: Hasmain, Seyed E.

; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa

; FILE REFERENCE: 6395-57049

; CURRENT FILING DATE: 2001-02-16

; PRIOR APPLICATION NUMBER: US 60/097,703

; PRIOR FILING DATE: 1998-08-21

; PRIOR APPLICATION NUMBER: PCT / US99/18869

; PRIOR FILING DATE: 1999-08-19

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 23

; LENGTH: 17

; TYPE: PRT

; ORGANISM: Plasmodium falciparum

US-09-763-397A-23

Query Match 100.0%; Score 84; DB 21; Length 17;  
Best Local Similarity 100.0%; Pred. No. 6.5e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LATRLMKFKAEIRDF 17

us-09-763-397a-23.rapm

Mon Feb 4 15:23:45 2002

```

RESULT 4
US-09-583-110-4656
; Sequence 4656, Application US/09583110
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococci
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4656
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-583-110-4656

Query Match 52.4%; Score 44; DB 19; Length 169;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 TRLMKKFKAEIRDF 16
Db 57 SRELKKYKHELQDF 70

RESULT 5
US-09-107-433-3442
; Sequence 3442, Application US/09107433
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; TITLE OF INVENTION: SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arieliello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3442:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
;

US-09-763-397A-2
; Sequence 2, Application US/09763397A
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis
; APPLICANT: Control and Prevention
; APPLICANT: Lal, Altaf A.
; APPLICANT: Lal, Altaf A.
; APPLICANT: Ping Shi, Ya
; APPLICANT: Hsainin, Seyed E.
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci
; FILE REFERENCE: 6395-57049
; CURRENT APPLICATION NUMBER: US/09/763,397A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/097,703
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: PCT / US99/18869
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant DNA/Protein
; US-09-763-397A-2

Query Match 100.0%; Score 84; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LATRLMKKFKAEIRDF 17
Db 317 LATRLMKKFKAEIRDF 333

RESULT 3
US-09-248-796-17817
; Sequence 17817, Application US/09248796
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 17817
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Candida albicans
; NAME/KEY: UNSURE
; LOCATION: (15)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
; US-09-248-796-17817

Query Match 53.6%; Score 45; DB 16; Length 743;
Best Local Similarity 64.3%; Pred. No. 2e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ATRLMKKFKAEIRDF 15
Db 242 ATRLKKOFNANITD 255

```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...174
; SEQUENCE DESCRIPTION: SEQ ID NO: 3442:
US-09-107-433-3442

Query Match      52.4%; Score 44; DB 15; Length 174;
Best Local Similarity 50.0%; Pred. No. 52;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 TRLMKKFAEIRDF 16
Db 62 SRFLKKYRHELODF 75

RESULT 6
US-60-269-308-4746
; Sequence 4746, Application US/60269308
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, RW
; APPLICANT: Ohlsen, KL
; APPLICANT: Zyskind, JW
; APPLICANT: Trawick, JD
; APPLICANT: Wall, D.
; TITLE OF INVENTION: Identification of Essential Genes in Staphylococcus aureus, Pseudomonas aeruginosa, Klebsiella pneumoniae, Salmonella typhimurium, and Haemophilus influenzae
; FILE REFERENCE: ELITRA 017PR5
; CURRENT APPLICATION NUMBER: US/60/269,308
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 4774
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4746
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-60-269-308-4746

Query Match      52.4%; Score 44; DB 24; Length 225;
Best Local Similarity 60.0%; Pred. No. 71;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 TRLMKKFAEIRDF 17
Db 92 TWTMRDFAEIRDEF 106

RESULT 7
US-09-417-507-25566
; Sequence 25566, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PATH99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 25566
; LENGTH: 62
; TYPE: PRT
; ORGANISM: A. fumigatus
US-09-417-507-25566

Query Match      51.2%; Score 43; DB 18; Length 62;
Best Local Similarity 51.2%; Score 43; DB 1; Length 86;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

US-09-763-397a-23.rapm

Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 TRLMKKFAEIR 13
Db 34 TRMKKEFKGEI 44

RESULT 8
PCT-US00-07726-121
; Sequence 121, Application PC/TUS00007726
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: PS538PCT
; CURRENT APPLICATION NUMBER: PCT/US00/07726
; CURRENT FILING DATE: 2000-03-23
; EARLIER APPLICATION NUMBER: 60/126,597
; EARLIER FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/174,877
; EARLIER FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 121
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-07726-121

Query Match      51.2%; Score 43; DB 1; Length 86;
Best Local Similarity 53.3%; Pred. No. 33;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LATRLMKKFAEIRD 15
Db 67 LATRFLKSEKANLEN 81

RESULT 9
PCT-US00-07726-122
; Sequence 122, Application PC/TUS00007726
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; FILE REFERENCE: PS538PCT
; CURRENT APPLICATION NUMBER: PCT/US00/07726
; CURRENT FILING DATE: 2000-03-23
; EARLIER APPLICATION NUMBER: 60/126,597
; EARLIER FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/174,877
; EARLIER FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 122
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-07726-122

Query Match      51.2%; Score 43; DB 1; Length 86;
Best Local Similarity 53.3%; Pred. No. 33;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LATRLMKKFAEIRD 15
Db 67 LATRFLKSEKANLEN 81

RESULT 10
PCT-US00-05988-1711
; Sequence 1711, Application PC/TUS0005988
```

```
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Human Prostate Cancer Associated Gene Sequences and Polypeptides
; FILE REFERENCE: PA101PCT
; CURRENT APPLICATION NUMBER: PCT/US00/05988
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/124,270
; EARLIER FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1711
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-05988-1711

Query Match 51.2%; Score 43; DB 1; Length 98;
Best Local Similarity 53.3%; Pred. No. 38;
Matches 8; Conservative 3; Mismatches 4; Indels 4; Gaps 0;

QY 1 LATRLMKKFKAEIRD 15
   |||| : ||| :
Db 79 LATRLKSKFKNLEN 93

RESULT 11
US-09-925-300-1711
; Sequence 1711, Application US/09925300
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1711
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1711

Query Match 51.2%; Score 43; DB 23; Length 98;
Best Local Similarity 53.3%; Pred. No. 38;
Matches 8; Conservative 3; Mismatches 4; Indels 4; Gaps 0;

QY 1 LATRLMKKFKAEIRD 15
   |||| : ||| :
Db 79 LATRLKSKFKNLEN 93

RESULT 12
US-09-146-394-1642
; Sequence 1642, Application US/60146394
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; FILE REFERENCE: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES THERE
; FILE REFERENCE: CL000063
; CURRENT APPLICATION NUMBER: US/60/146,394
; CURRENT FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 1866
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1642
; LENGTH: 155
```

```
; TYPE: PRT
; ORGANISM: Drosophila
US-60-146-394-1642
```

```
Query Match 51.2%; Score 43; DB 24; Length 155;
Best Local Similarity 81.8%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 LATRLMKKFKKA 11
   ||| : |||||
Db 53 LALRVKKFKKA 63
```

```
RESULT 13
US-09-760-475-1793
; Sequence 1793, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ49
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1793
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-475-1793
```

```
Query Match 51.2%; Score 43; DB 21; Length 301;
Best Local Similarity 53.3%; Pred. No. 1.5e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 LATRLMKKFKAEIRD 15
   |||| : ||| :
Db 282 LATRLKSKFKNLEN 296
```

```
RESULT 14
US-09-760-475-3012
; Sequence 3012, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ49
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3012
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-475-3012
```

```
Query Match 51.2%; Score 43; DB 21; Length 316;
Best Local Similarity 53.3%; Pred. No. 1.6e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 LATRLMKKFKAEIRD 15
   |||| : ||| :
Db 297 LATRLKSKFKNLEN 311
```

```
RESULT 15
US-60-245-201-221
; Sequence 221, Application US/60245201
```

; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PHASE II  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: HUMAN DRUG-METABOLIZING PHASE II PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CLO00879  
; CURRENT APPLICATION NUMBER: US/60/245,201  
; CURRENT FILING DATE: 2000-11-03  
; NUMBER OF SEQ ID NOS: 381  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 221  
; LENGTH: 509  
; TYPE: PRT  
; ORGANISM: Human  
US-60-245-201-221

Query Match 51.2%; Score 43; DB 24; Length 509;  
Best Local Similarity 53.3%; Pred. No. 2.8e+02;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LATRLMKKEKAIRD 15  
|||:| |||:  
Db 490 LATRLKSPKANLEN 504

Search completed: January 29, 2002, 11:06:40  
Job time: 622 sec

**THIS PAGE BLANK (USPTO)**



```
RESULT 2
US-09-815-242-10689
; Sequence 10689, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.Oila
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10689
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10689

Query Match 52.48; Score 44; DB 5; Length 225;
Best Local Similarity 60.08; Pred. No. 5.4;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 TRLMKKFAEIRDF 17
| | | | | | | |
DB 92 TNTMRDFAEIRDEF 106

RESULT 3
US-09-708-427-10555
; Sequence 10555, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10555
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..297
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..297
; OTHER INFORMATION: Ceres Seq. ID 1821590
US-09-708-427-10555
```

```
Query Match 51.28; Score 43; DB 5; Length 297;
Best Local Similarity 58.38; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 MKKFKAEIRDF 17
: |||:|:|
DB 36 LDKFKSDIQDF 47

RESULT 4
US-09-708-427-10554
; Sequence 10554, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10554
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..335
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..335
; OTHER INFORMATION: Ceres Seq. ID 1821589
US-09-708-427-10554
```

```
Query Match 51.28; Score 43; DB 5; Length 335;
Best Local Similarity 58.38; Pred. No. 12;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 MKKFKAEIRDF 17
: |||:|:|
DB 74 LDKFKSDIQDF 85

RESULT 5
US-09-708-427-10553
; Sequence 10553, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10553
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..353
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..353
; OTHER INFORMATION: Ceres Seq. ID 1821588
US-09-708-427-10553
```

Query Match 51.28; Score 43; DB 5; Length 353;



Best Local Similarity 58.3%; Pred. No. 13;  
Matches 7; Conservative 4; Mismatches

1; Indels 0; Gaps 0;

QY 6 MKKFAEIRDF 17

Db 92 LDKFKSDIQDF 103

RESULT 6

US-09-708-427-30719  
; Sequence 30719, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 30719  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..313  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..313  
; OTHER INFORMATION: Ceres Seq. ID 1830099  
US-09-708-427-30719

Query Match 50.0%; Score 42; DB 5; Length 313;

Best Local Similarity 60.0%; Pred. No. 17;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATRLMKKFAEIRDF 16

Db 80 ATRLFKKLVEIMF 94

RESULT 7

US-09-708-427-30718  
; Sequence 30718, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 30718  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..318  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..318  
; OTHER INFORMATION: Ceres Seq. ID 1830098  
US-09-708-427-30718

Query Match 50.0%; Score 42; DB 5; Length 318;

Best Local Similarity 60.0%; Pred. No. 17;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATRLMKKFAEIRDF 16

Db 85 ATRLFKKLVEIMF 99

RESULT 8

US-09-708-427-30717  
; Sequence 30717, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 30717  
; LENGTH: 398  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..398  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..398  
; OTHER INFORMATION: Ceres Seq. ID 1830097  
US-09-708-427-30717

Query Match 50.0%; Score 42; DB 5; Length 398;

Best Local Similarity 60.0%; Pred. No. 22;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATRLMKKFAEIRDF 16

Db 165 ATRLFKKLVEIMF 179

RESULT 9

US-09-708-427-29820  
; Sequence 29820, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 29820  
; LENGTH: 593  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..593  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..593  
; OTHER INFORMATION: Ceres Seq. ID 1827158  
US-09-708-427-29820

Query Match 50.0%; Score 42; DB 5; Length 593;

Best Local Similarity 61.5%; Pred. No. 34;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 LMKKFAEIRDF 17

Db 135 LMKKFQSTRDLF 147

```
RESULT 10
US-09-708-427-29819
; Sequence 29819, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 29819
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..686
; OTHER INFORMATION: xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..686
; OTHER INFORMATION: Ceres Seq. ID 1827157
US-09-708-427-29819

Query Match          50.0%; Score 42; DB 5; Length 686;
Best Local Similarity 61.5%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 4; Indels 4; Gaps 0;

QY 5 LMKKFKAEIRDF 17
DB 228 LMKKFTSTRDLF 240

RESULT 11
US-09-708-427-29818
; Sequence 29818, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 29818
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..743
; OTHER INFORMATION: xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..743
; OTHER INFORMATION: Ceres Seq. ID 1827156
US-09-708-427-29818

Query Match          50.0%; Score 42; DB 5; Length 743;
Best Local Similarity 61.5%; Pred. No. 44;
Matches 8; Conservative 1; Mismatches 4; Indels 4; Gaps 0;

QY 5 LMKKFKAEIRDF 17
DB 285 LMKKFTSTRDLF 297

RESULT 12
US-09-708-427-19349
; Sequence 19349, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19349
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..303
; OTHER INFORMATION: xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..303
; OTHER INFORMATION: Ceres Seq. ID 1835886
US-09-708-427-19349

Query Match          48.8%; Score 41; DB 5; Length 303;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 RLMKKFKAEIRD 15
DB 128 RVMKKVQAEIRE 139

RESULT 13
US-09-708-427-19348
; Sequence 19348, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19348
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..314
; OTHER INFORMATION: xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..314
; OTHER INFORMATION: Ceres Seq. ID 1835885
US-09-708-427-19348

Query Match          48.8%; Score 41; DB 5; Length 314;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 RLMKKFKAEIRD 15
DB 139 RVMKKVQAEIRE 150

RESULT 14
US-09-708-427-19347
; Sequence 19347, Application US/09708427
; GENERAL INFORMATION:
```

; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19347  
; LENGTH: 503  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..503  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..503  
; OTHER INFORMATION: Ceres Seq. ID 1835884  
US-09-708-427-19347

Query Match 48.8%; Score 41; DB 5; Length 503;  
Best Local Similarity 66.7%; Pred. No. 41;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RLKKKFAEIRD 15  
I:|:|:|:|:|:|:  
Db 328 RVMKKVQAEIRE 339

RESULT 15  
US-09-708-427-58150  
; Sequence 58150, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 58150  
; LENGTH: 302  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..302  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..302  
; OTHER INFORMATION: Ceres Seq. ID 1941286  
US-09-708-427-58150

Query Match 47.6%; Score 40; DB 5; Length 302;  
Best Local Similarity 53.8%; Pred. No. 34;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RLKKKFAEIRDF 16  
I:|:|:|:|:|:|:  
Db 119 RVMKKLQAEVRAY 131

Search completed: January 29, 2002, 11:07:15  
Job time: 533 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:18 ; Search time 353.23 seconds  
(without alignments)  
13.363 Million cell updates/sec

Title: US-09-763-397A-22

Perfect score: 79

Sequence: 1 SSPSTKSPSNVKSAS 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	17	21	US-09-763-397A-22
2	68.5	86.7	350	21	US-09-763-397A-2
3	50	63.3	2559	24	US-60-191-637-15071
4	50	63.3	2559	24	US-60-191-681-11903
5	49	62.0	81	21	US-09-758-443-408
6	48	60.8	179	24	US-60-324-109-26094
7	48	60.8	328	1	PCT-US00-31739-10
8	48	60.8	328	1	PCT-US99-08954-2
9	48	60.8	328	21	US-09-715-774-10

10	46	58.2	106	1	PCT-US01-08631-37162
11	46	58.2	152	24	US-60-324-109-18117
12	46	58.2	174	24	US-60-169-840-5493
13	46	58.2	176	24	US-60-177-571-3322
14	46	58.2	198	21	US-09-758-459-219
15	46	58.2	210	24	US-60-173-469-1145
16	46	58.2	243	1	PCT-US00-26337-78
17	46	58.2	273	17	US-09-396-149-12
18	46	58.2	273	17	US-09-396-149-14
19	46	58.2	273	17	US-09-396-149-16
20	46	58.2	273	17	US-09-396-149-20
21	46	58.2	273	17	US-09-396-149-22
22	46	58.2	303	18	US-09-488-725A-6974
23	46	58.2	310	1	PCT-US00-14973-60
24	46	58.2	310	1	PCT-US00-14973-110
25	46	58.2	310	1	PCT-US00-26376-116
26	46	58.2	310	1	PCT-US01-16450-2703
27	46	58.2	310	1	PCT-US01-16450A-2703
28	46	58.2	310	18	US-09-488-725A-3402
29	46	58.2	310	23	US-09-950-083-3186
30	46	58.2	310	23	US-09-950-083-4481
31	46	58.2	311	1	PCT-US00-26376-74
32	46	58.2	318	17	US-09-396-149-18
33	46	58.2	462	16	US-09-248-796-14808
34	45	57.0	114	1	PCT-US01-08631-33984
35	45	57.0	205	1	PCT-US00-26524B-6129
36	45	57.0	241	1	PCT-US01-08631-33985
37	45	57.0	362	16	US-09-252-991A-17792
38	45	57.0	483	24	US-60-312-344-8278
39	45	57.0	625	20	US-09-618-425-13
40	45	57.0	626	13	US-08-971-188-10
41	45	57.0	626	17	US-09-374-454-21
42	45	57.0	2274	24	US-60-161-932-1800
43	45	57.0	2380	24	US-60-167-217-19522
44	45	57.0	2380	24	US-60-173-464-15956
45	45	57.0	2406	24	US-60-191-681-15359

ALIGNMENTS

RESULT 1  
US-09-763-397A-22  
; Sequence 22, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-22

Query Match 100.0% Score 79; DB 21; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SSPSTKSPSNVKSAS 17

```

; Sequence 11903, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USSES THEREOF.
; FILE REFERENCE: C1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11903
; LENGTH: 2559
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-11903

Query Match 63.3%; Score 50; DB 24; Length 2559;
Best Local Similarity 76.9%; Pred. No. 7.9e+02;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PSSTKSSPSNVKS 15
Db 645 PSSTNDSPSNAKS 657

RESULT 5
US-09-758-443-408
; Sequence 408, Application US/09758443
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM039
; CURRENT APPLICATION NUMBER: US/09/758,443
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 432
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 408
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (51)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-443-408

Query Match 62.0%; Score 49; DB 21; Length 81;
Best Local Similarity 58.8%; Pred. No. 21;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSPSTKSSPSNVKSAS 17
Db 23 SSPTTKSSPSDPMTTS 39

RESULT 6
US-60-324-109-26094

```

```

; Sequence 11903, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USSES THEREOF.
; FILE REFERENCE: C1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11903
; LENGTH: 2559
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-11903

Query Match 63.3%; Score 50; DB 24; Length 2559;
Best Local Similarity 76.9%; Pred. No. 7.9e+02;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PSSTKSSPSNVKS 15
Db 645 PSSTNDSPSNAKS 657

RESULT 5
US-09-758-443-408
; Sequence 408, Application US/09758443
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM039
; CURRENT APPLICATION NUMBER: US/09/758,443
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 432
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 408
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (51)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-443-408

Query Match 62.0%; Score 49; DB 21; Length 81;
Best Local Similarity 58.8%; Pred. No. 21;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSPSTKSSPSNVKSAS 17
Db 23 SSPTTKSSPSDPMTTS 39

RESULT 6
US-60-324-109-26094

```

```

; Sequence 11903, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USSES THEREOF.
; FILE REFERENCE: C1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11903
; LENGTH: 2559
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-11903

Query Match 63.3%; Score 50; DB 24; Length 2559;
Best Local Similarity 76.9%; Pred. No. 7.9e+02;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PSSTKSSPSNVKS 15
Db 645 PSSTNDSPSNAKS 657

RESULT 5
US-09-758-443-408
; Sequence 408, Application US/09758443
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM039
; CURRENT APPLICATION NUMBER: US/09/758,443
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 432
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant DNA/Protein
US-09-763-397A-2

Query Match 86.7%; Score 68.5; DB 21; Length 350;
Best Local Similarity 94.4%; Pred. No. 0.16;
Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 SSPSTKSSPSNVKSAS 17
Db 299 SSPSTKSSPSNVKSAS 316

RESULT 3
US-60-191-637-15071
; Sequence 15071, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USSES THEREOF
; FILE REFERENCE: C1000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15071
; LENGTH: 2559
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-15071

Query Match 63.3%; Score 50; DB 24; Length 2559;
Best Local Similarity 76.9%; Pred. No. 7.9e+02;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PSSTKSSPSNVKS 15
Db 645 PSSTNDSPSNAKS 657

RESULT 4
US-60-191-681-11903

```

; Sequence 26094, Application US/60324109  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Stein, Joshua  
; TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-1015272618  
; CURRENT APPLICATION NUMBER: US/60/324,109  
; CURRENT FILING DATE: 2001-09-21  
; NUMBER OF SEQ ID NOS: 33196  
; SEQ ID NO 26094  
; LENGTH: 179  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; US-60-324-109-26094

Query Match 60.8%; Score 48; DB 24; Length 179;  
Best Local Similarity 64.7%; Pred. No. 74;  
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSPSSTKSSPSNVKAS 17  
||||| ||| : |||  
Db 144 SSPASSSSSSSSSAS 160

RESULT 7  
PCT-US00-31739-10  
; Sequence 10, Application PC/TUS0031739  
; GENERAL INFORMATION:  
; APPLICANT: Helentjaris, Tim  
; TITLE OF INVENTION: Modulation of Absciscic Acid  
; FILE REFERENCE: 35718/205302  
; CURRENT APPLICATION NUMBER: PCT/US00/31739  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: US 60/166,080  
; PRIOR FILING DATE: 1999-11-17  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 328  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; PCT-US00-31739-10

Query Match 60.8%; Score 48; DB 1; Length 328;  
Best Local Similarity 64.7%; Pred. No. 1.5e+02;  
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSPSSTKSSPSNVKAS 17  
||||| ||| : |||  
Db 113 SSPSSVSSSSSVSAAS 129

RESULT 8  
PCT-US99-08954-2  
; Sequence 2, Application PC/TUS9908954  
; GENERAL INFORMATION:  
; APPLICANT: Finkelstein, Ruth R.  
; APPLICANT: Lynch, Tim  
; APPLICANT: Goodman, Howard M.  
; APPLICANT: Wang, Ming-Li  
; TITLE OF INVENTION: A TRANSCRIPTION FACTOR REGULATING SEED DEVELOPMENT,  
; FILE REFERENCE: 480.89(HV)  
; CURRENT APPLICATION NUMBER: PCT/US99/08954  
; CURRENT FILING DATE: 1999-04-27  
; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 328  
; TYPE: PRT  
; ORGANISM: Arabidopsis  
; PCT-US99-08954-2

Query Match 60.8%; Score 48; DB 1; Length 328;  
Best Local Similarity 64.7%; Pred. No. 1.5e+02;  
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSPSSTKSSPSNVKAS 17  
||||| ||| : |||  
Db 113 SSPSSVSSSSSVSAAS 129

RESULT 9  
US-09-715-774-10  
; Sequence 10, Application US/09715774  
; GENERAL INFORMATION:  
; APPLICANT: Helentjaris, Tim  
; TITLE OF INVENTION: Modulation of Absciscic Acid  
; FILE REFERENCE: 35718/204777  
; CURRENT APPLICATION NUMBER: US/09/715,774  
; CURRENT FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: US 60/166,080  
; PRIOR FILING DATE: 1999-11-17  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 328  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; US-09-715-774-10

Query Match 60.8%; Score 48; DB 21; Length 328;  
Best Local Similarity 64.7%; Pred. No. 1.5e+02;  
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSPSSTKSSPSNVKAS 17  
||||| ||| : |||  
Db 113 SSPSSVSSSSSVSAAS 129

RESULT 10  
PCT-US01-08631-37162  
; Sequence 37162, Application PC/TUS0108631  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-049  
; CURRENT APPLICATION NUMBER: PCT/US01/08631  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 37162  
; LENGTH: 106  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (34)...(83)  
; OTHER INFORMATION: Eukaryotic RNA polymerase II heptapeptide repeat proteins  
; OTHER INFORMATION: domain identified by eMATRIX, accession number BL001152, p-val  
; OTHER INFORMATION: 5.103e-09, raw score of 3.12  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(106)

OTHER INFORMATION: Xaa = X or \* as defined in Table 2  
PCT-US01-08631-37162

Query Match 58.2%; Score 46; DB 1; Length 106;  
Best Local Similarity 58.8%; Pred. No. 79;  
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSPSTKSSPSNVKSAS 17  
||||| :|| :|||  
Db 41 SSPSSSPSPSSSSSS 57

RESULT 11  
US-60-324-109-18117  
; Sequence 18117, Application US/60324109  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Stein, Joshua  
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-10(52726)B  
; CURRENT APPLICATION NUMBER: US/60/324,109  
; CURRENT FILING DATE: 2001-09-21  
; NUMBER OF SEQ ID NOS: 33196  
; SEQ ID NO 18117  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
US-60-324-109-18117

Query Match 58.2%; Score 46; DB 24; Length 152;  
Best Local Similarity 58.8%; Pred. No. 1.2e+02;  
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSPSTKSSPSNVKSAS 17  
|:|:| ||||: |||  
Db 69 STPTSLKSPAPVTS 85

RESULT 12  
US-60-169-840-5493  
; Sequence 5493, Application US/60169840  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CLO00164  
; CURRENT APPLICATION NUMBER: US/60/169,840  
; CURRENT FILING DATE: 1999-12-09  
; NUMBER OF SEQ ID NOS: 9628  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5493  
; LENGTH: 174  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(174)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-60-169-840-5493

Query Match 58.2%; Score 46; DB 24; Length 174;  
Best Local Similarity 64.7%; Pred. No. 1.4e+02;  
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSPSTKSSPSNVKSAS 17  
||||| :|| :|||  
Db 145 SSPSSLSTSPPEVSAS 161

RESULT 13  
US-60-177-571-3322  
; Sequence 3322, Application US/60177571  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CLO00201  
; CURRENT APPLICATION NUMBER: US/60/177,571  
; CURRENT FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 5082  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3322  
; LENGTH: 176  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-177-571-3322

Query Match 58.2%; Score 46; DB 24; Length 176;  
Best Local Similarity 64.7%; Pred. No. 1.4e+02;  
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSPSTKSSPSNVKSAS 17  
||||| :|| :|||  
Db 100 SSPSSLSTSPPEVSAS 116

RESULT 14  
US-09-758-459-219  
; Sequence 219, Application US/09758459  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PM029  
; CURRENT APPLICATION NUMBER: US/09/758,459  
; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; NUMBER OF SEQ ID NOS: 406  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 219  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-459-219

Query Match 58.2%; Score 46; DB 21; Length 198;  
Best Local Similarity 64.7%; Pred. No. 1.6e+02;  
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSPSTKSSPSNVKSAS 17  
||||| :|| :|||  
Db 126 SSPSSLSTSPPEVSAS 142

RESULT 15  
US-60-173-469-1145  
; Sequence 1145, Application US/60173469  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND  
; TITLE OF INVENTION: USES THEREOF



; FILE REFERENCE: CL000185  
; CURRENT APPLICATION NUMBER: US/60/173.469  
; CURRENT FILING DATE: 1999-12-29  
; NUMBER OF SEQ ID NOS: 2120  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1145  
; LENGTH: 210  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-173-469-1145

Query Match 58.2%; Score 46; DB 24; Length 210;  
Best Local Similarity 64.7%; Pred. NO. 1.7e+02;  
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSPSTKSSPSNVKSAS 17  
|||||..||| |||  
Db 135 SSPSLSTSPPEVFSAS 151

Search completed: January 29, 2002, 11:06:39  
Job time: 621 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:22 ; Search time 29.91 Seconds  
(without alignments)  
20.801 Million cell updates/sec

Title: US-09-763-397a-22

Perfect score: 79  
Sequence: 1 SSPSTKSSPSNVKSAS 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	48	60.8	328	5	US-09-708-427-16088
2	47	59.5	515	5	US-09-801-368-384
3	47	59.5	515	5	US-09-487-558-384
4	46	58.2	200	5	US-09-637-780B-120
5	46	58.2	208	5	US-09-637-780B-119
6	46	58.2	272	5	US-09-708-427-82455
7	46	58.2	273	5	US-09-708-427-82454
8	46	58.2	297	5	US-09-637-780B-118
9	46	58.2	297	5	US-09-708-427-82453
10	44	55.7	229	5	US-09-708-427-11927
11	44	55.7	1157	5	US-09-708-427-3137
12	44	55.7	1170	5	US-09-708-427-3136
13	44	55.7	1194	5	US-09-708-427-3135
14	43	54.4	215	5	US-09-620-394B-3870
15	43	54.4	231	5	US-09-620-394B-3869
16	43	54.4	323	5	US-09-708-427-3689
17	43	54.4	365	5	US-09-708-427-3688
18	43	54.4	402	5	US-09-708-427-3687
19	43	54.4	490	5	US-09-708-427-33009
20	43	54.4	537	5	US-09-708-427-33008
21	43	54.4	541	5	US-09-708-427-33007
22	43	54.4	542	5	US-09-708-427-5389
23	43	54.4	969	5	US-09-708-427-5596
24	43	54.4	969	5	US-09-708-427-76259
25	43	54.4	992	5	US-09-708-427-76258
26	43	54.4	1036	6	US-10-005-368-205

27	42	53.2	81	5	US-09-620-111B-4982	Sequence 4982, Ap
28	42	53.2	118	5	US-09-620-111B-4981	Sequence 4981, Ap
29	42	53.2	283	5	US-09-708-427-3568	Sequence 3568, Ap
30	42	53.2	362	5	US-09-708-427-17129	Sequence 17129, A
31	42	53.2	363	5	US-09-708-427-17128	Sequence 17128, A
32	42	53.2	380	5	US-09-708-427-17127	Sequence 17127, A
33	42	53.2	496	5	US-09-708-427-3567	Sequence 3567, Ap
34	42	53.2	503	5	US-09-801-368-424	Sequence 424, App
35	42	53.2	503	5	US-09-487-558-424	Sequence 424, App
36	42	53.2	944	5	US-09-964-238-2	Sequence 2, Appli
37	41	51.9	104	5	US-09-620-111B-4152	Sequence 4152, Ap
38	41	51.9	111	5	US-09-708-427-47194	Sequence 47194, A
39	41	51.9	125	5	US-09-620-111B-761	Sequence 761, App
40	41	51.9	135	5	US-09-708-427-47193	Sequence 47193, A
41	41	51.9	145	5	US-09-620-111B-4151	Sequence 4151, Ap
42	41	51.9	153	5	US-09-620-111B-760	Sequence 760, App
43	41	51.9	198	5	US-09-708-427-66585	Sequence 66585, A
44	41	51.9	2016	5	US-09-897-516-7376	Sequence 7376, Ap
45	41	51.9	6025	6	US-10-005-368-61	Sequence 61, Appli

ALIGNMENTS

RESULT 1

US-09-708-427-16088  
; Sequence 16088, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16088  
; LENGTH: 328  
; TYPE: PPT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..328  
; OTHER INFORMATION: xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..328  
; OTHER INFORMATION: Ceres Seq. ID 1830172  
US-09-708-427-16088

Query Match 60.8%; Score 48; DB 5; Length 328;

Best Local Similarity 64.7%; Pred. No. 2.7; Mismatches 2; Indels 0; Gaps 0;

Matches 11; Conservative 4; Indels 0; Gaps 0;

Qy 1 SSPSTKSSPSNVKSAS 17

Db 113 SSPSSVSSSSSVSAAS 129

RESULT 2

US-09-801-368-384  
; Sequence 384, Application US/09801368  
; GENERAL INFORMATION:  
; APPLICANT: Busby, Robert  
; APPLICANT: Cali, Brian  
; APPLICANT: Hecht, Peter  
; APPLICANT: Holtzman, Doug  
; APPLICANT: Madden, Kevin  
; APPLICANT: Maxon, Mary  
; APPLICANT: Milne, Todd  
; APPLICANT: Norman, Thea  
; APPLICANT: Royer, John  
; APPLICANT: Salama, Sofie

us-09-763-397a-22.rapn

Mon Feb 4 15:23:44 2002

```

; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 384
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-801-368-384

```

Query Match 59.5%; Score 47; DB 5; Length 515;  
 Best Local Similarity 56.2%; Pred. No. 6.3;  
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 SPSTKSSPSNVKAS 17  
 ||||| ||||| :|  
 Db 130 SPSTNSTPSTIQGLS 145

RESULT 3  
 US-09-487-558-384  
 ; Sequence 384, Application US/09487558  
 ; GENERAL INFORMATION:

```

; APPLICANT: Busby, Robert
; APPLICANT: Calli, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Norman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/487,558
; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US/09/801,368
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 384
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-487-558-384

```

Query Match 59.5%; Score 47; DB 5; Length 515;  
 Best Local Similarity 56.2%; Pred. No. 6.3;  
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 SPSTKSSPSNVKAS 17  
 ||||| ||||| :|  
 Db 130 SPSTNSTPSTIQGLS 145

```

RESULT 4
US-09-637-780B-120
; Sequence 120, Application US/09637780B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1096P
; CURRENT APPLICATION NUMBER: US/09/637,780B
; CURRENT FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 1401
; SEQ ID NO 120
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..200
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..200
; OTHER INFORMATION: Ceres Seq. ID 1481703
; US-09-637-780B-120

```

Query Match 58.2%; Score 46; DB 5; Length 200;  
 Best Local Similarity 58.8%; Pred. No. 3.3;  
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SPSTKSSPSNVKAS 17  
 |::| ||||: |||  
 Db 117 STPTSLKSPAPVTS GS 133

```

RESULT 5
US-09-637-780B-119
; Sequence 119, Application US/09637780B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1096P
; CURRENT APPLICATION NUMBER: US/09/637,780B
; CURRENT FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 1401
; SEQ ID NO 119
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..208
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..208
; OTHER INFORMATION: Ceres Seq. ID 1481702
; US-09-637-780B-119

```

Query Match 58.2%; Score 46; DB 5; Length 208;  
 Best Local Similarity 58.8%; Pred. No. 3.4;  
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SPSTKSSPSNVKAS 17  
 |::| ||||: |||  
 Db 125 STPTSLKSPAPVTS GS 141

RESULT 6

US-09-708-427-82455  
; Sequence 82455, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 82455  
; LENGTH: 272  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..272  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..272  
; OTHER INFORMATION: Ceres Seq. ID 1960819  
US-09-708-427-82455

Query Match 58.2%; Score 46; DB 5; Length 272;  
Best Local Similarity 58.8%; Pred. No. 4.5;  
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSPSSTKSSPSNVKAS 17  
I:|:| | | | | | | | | |  
Db 189 STPTSLKSSPAPVTSGS 205

RESULT 7  
US-09-708-427-82454  
; Sequence 82454, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 82454  
; LENGTH: 273  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..273  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..273  
; OTHER INFORMATION: Ceres Seq. ID 1960818  
US-09-708-427-82454

Query Match 58.2%; Score 46; DB 5; Length 273;  
Best Local Similarity 58.8%; Pred. No. 4.6;  
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSPSSTKSSPSNVKAS 17  
I:|:| | | | | | | | | |  
Db 190 STPTSLKSSPAPVTSGS 206

RESULT 8  
US-09-637-780B-118  
; Sequence 118, Application US/09637780B  
; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nickolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1096P  
; CURRENT APPLICATION NUMBER: US/09/637,780B  
; CURRENT FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 1401  
; SEQ ID NO 118  
; LENGTH: 297  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..297  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..297  
; OTHER INFORMATION: Ceres Seq. ID 1481701  
US-09-637-780B-118

Query Match 58.2%; Score 46; DB 5; Length 297;  
Best Local Similarity 58.8%; Pred. No. 5;  
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSPSSTKSSPSNVKAS 17  
I:|:| | | | | | | | | |  
Db 214 STPTSLKSSPAPVTSGS 230

RESULT 9  
US-09-708-427-82453  
; Sequence 82453, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 82453  
; LENGTH: 297  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..297  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..297  
; OTHER INFORMATION: Ceres Seq. ID 1960817  
US-09-708-427-82453

Query Match 58.2%; Score 46; DB 5; Length 297;  
Best Local Similarity 58.8%; Pred. No. 5;  
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSPSSTKSSPSNVKAS 17  
I:|:| | | | | | | | | |  
Db 214 STPTSLKSSPAPVTSGS 230

RESULT 10  
US-09-708-427-11927  
; Sequence 11927, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; TITLE OF INVENTION: THEREBY

```
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11927
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..229
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..229
; OTHER INFORMATION: Ceres Seq. ID 1824027
; US-09-708-427-11927

Query Match          55.7%; Score 44; DB 5; Length 229;
Best Local Similarity 58.8%; Pred. No. 7.7;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSPSSTKSSPSNVKAS 17
   :||||: ||| :||
Db 35 SAPSSSSSSSTVAAS 51

RESULT 11
US-09-708-427-3137
; Sequence 3137, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3137
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1157
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1157
; OTHER INFORMATION: Ceres Seq. ID 1810132
; US-09-708-427-3137

Query Match          55.7%; Score 44; DB 5; Length 1157;
Best Local Similarity 52.9%; Pred. No. 44;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSPSSTKSSPSNVKAS 17
   :||||: ||| :||
Db 801 NSPSSSSSSPTPIPS 817

RESULT 12
US-09-708-427-3136
; Sequence 3136, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
```

```
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3136
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1170
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1170
; OTHER INFORMATION: Ceres Seq. ID 1810131
; US-09-708-427-3136

Query Match          55.7%; Score 44; DB 5; Length 1170;
Best Local Similarity 52.9%; Pred. No. 45;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSPSSTKSSPSNVKAS 17
   :||||: ||| :||
Db 814 NSPSSSSSSPTPIPS 830

RESULT 13
US-09-708-427-3135
; Sequence 3135, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3135
; LENGTH: 1194
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1194
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1194
; OTHER INFORMATION: Ceres Seq. ID 1810130
; US-09-708-427-3135

Query Match          55.7%; Score 44; DB 5; Length 1194;
Best Local Similarity 52.9%; Pred. No. 46;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSPSSTKSSPSNVKAS 17
   :||||: ||| :||
Db 838 NSPSSSSSSPTPIPS 854

RESULT 14
US-09-620-394B-3870
; Sequence 3870, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 3870
```

```

; LENGTH: 215
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..215
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..215
; OTHER INFORMATION: Ceres Seq. ID 1387751
US-09-620-394B-3870

Query Match          54.4%; Score 43; DB 5; Length 215;
Best Local Similarity 58.8%; Pred. No. 10;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSPSSTKSSPSNVKAS 17
Db 5 SSPSSSSSSSSSSSS 21

RESULT 15
US-09-620-394B-3869
; Sequence 3869, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 3869
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..231
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..231
; OTHER INFORMATION: Ceres Seq. ID 1387750
US-09-620-394B-3869

Query Match          54.4%; Score 43; DB 5; Length 231;
Best Local Similarity 58.8%; Pred. No. 11;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSPSSTKSSPSNVKAS 17
Db 21 SSPSSSSSSSSSSSS 37

Search completed: January 29, 2002, 11:07:15
Job time: 533 sec
```

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:13 ; Search time 1760.55 Seconds  
(without alignments)  
1.262 Million cell updates/sec

Title: US-09-763-397A-21  
Perfect score: 41  
Sequence: 1 LTPLEELY 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main: \*  
1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	41	100.0	8	21	US-09-763-397A-21
2	41	100.0	9	15	US-09-185-878-60
3	41	100.0	350	21	US-09-763-397A-2
4	39	95.1	242	21	US-09-760-475-2603
5	39	95.1	493	23	US-09-948-947-113
6	37	90.2	88	18	US-09-417-507-35531
7	37	90.2	310	24	US-60-316-362-3517
8	37	90.2	338	19	US-09-543-681A-4430
9	36	87.8	284	1	PCT-US00-26524B-4656

10	36	87.8	806	17	US-09-328-352-7160	Sequence 7160, Ap
11	35	85.4	71	24	US-60-146-055-775	Sequence 775, App
12	35	85.4	71	24	US-60-147-189-849	Sequence 849, App
13	35	85.4	148	16	US-09-270-767-60152	Sequence 60152, A
14	35	85.4	148	16	US-09-270-849B-189381	Sequence 189381, A
15	35	85.4	329	16	US-09-270-767-44698	Sequence 44698, A
16	35	85.4	422	24	US-60-173-464-7042	Sequence 7042, Ap
17	35	85.4	443	24	US-60-191-637-8717	Sequence 8717, Ap
18	35	85.4	443	24	US-60-191-681-6776	Sequence 6776, Ap
19	35	85.4	445	16	US-09-270-767-45644	Sequence 45644, A
20	34	82.9	283	16	US-09-201-228A-181	Sequence 181, App
21	34	82.9	414	22	US-09-841-132-568	Sequence 568, App
22	33	80.5	155	18	US-09-488-725A-2441	Sequence 2441, Ap
23	33	80.5	305	15	US-09-107-532-5749	Sequence 5749, Ap
24	33	80.5	305	15	US-09-107-532A-5749	Sequence 5749, Ap
25	33	80.5	398	12	US-08-827-356-4601	Sequence 4601, Ap
26	33	80.5	398	20	US-09-611-529-7191	Sequence 7191, Ap
27	33	80.5	406	15	US-09-134-000-6011	Sequence 6011, Ap
28	33	80.5	506	24	US-60-167-217-22857	Sequence 22857, A
29	33	80.5	506	24	US-60-173-386-1993	Sequence 1993, Ap
30	33	80.5	506	24	US-60-175-871-2230	Sequence 2230, Ap
31	33	80.5	506	24	US-60-184-775-2041	Sequence 2041, Ap
32	33	80.5	506	24	US-60-191-637-42612	Sequence 42612, A
33	33	80.5	506	24	US-60-191-700-2169	Sequence 2169, Ap
34	33	80.5	534	18	US-09-417-507-37851	Sequence 37851, A
35	33	80.5	673	1	PCT-US01-14827-9961	Sequence 9961, App
36	33	80.5	691	24	US-60-150-584-964	Sequence 964, App
37	33	80.5	743	24	US-60-259-128-4681	Sequence 4681, Ap
38	33	80.5	743	24	US-60-314-050-7145	Sequence 7145, Ap
39	33	80.5	1230	24	US-60-316-362-3281	Sequence 3281, Ap
40	32	78.0	76	18	US-09-488-725A-2730	Sequence 2730, Ap
41	32	78.0	85	18	US-09-488-725A-6302	Sequence 6302, Ap
42	32	78.0	177	16	US-09-270-767-32992	Sequence 32992, A
43	32	78.0	304	1	PCT-US01-18342-11	Sequence 11, Appl
44	32	78.0	304	19	US-09-587-789-11	Sequence 11, Appl
45	32	78.0	327	24	US-60-207-315-308	Sequence 308, App

ALIGNMENTS

RESULT 1  
US-09-763-397A-21  
; Sequence 21, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 21  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-21

Query Match 100.0%; Score 41; DB 21; Length 8;  
Best Local Similarity 100.0%; Pred No. 2.9e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LTPLEELY 8

Db 1 LTPLEELY 8  
|||||

RESULT 2  
US-09-165-878-60  
; Sequence 60, Application US/09165878B  
; GENERAL INFORMATION:  
; APPLICANT: Heegaard, Peter Mikael Helweg  
; APPLICANT: Jakobsen, Palle Hoy  
; TITLE OF INVENTION: Non-Dendritic Backbone Peptide Carrier  
; FILE REFERENCE: 2316.1009-000  
; CURRENT APPLICATION NUMBER: US/09/165,878B  
; CURRENT FILING DATE: 1998-10-02  
; EARLIER APPLICATION NUMBER: PCT/DK97/00146  
; EARLIER FILING DATE: 1997-04-03  
; EARLIER APPLICATION NUMBER: DK 0398/96  
; EARLIER FILING DATE: 1996-04-03  
; NUMBER OF SEQ ID NOS: 113  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 60  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Unknown  
US-09-165-878-60

Query Match 100.0%; Score 41; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTPLEELY 8  
|||||

Db 1 LTPLEELY 8  
|||||

RESULT 3  
US-09-763-397A-2  
; Sequence 2, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Recombinant DNA/Protein  
US-09-763-397A-2

Query Match 100.0%; Score 41; DB 21; Length 350;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTPLEELY 8  
|||||

Db 176 LTPLEELY 183

RESULT 4  
US-09-760-475-2603  
; Sequence 2603, Application US/09760475  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT249  
; CURRENT APPLICATION NUMBER: US/09/760,475  
; CURRENT FILING DATE: 2001-01-16  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4122  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2603  
; LENGTH: 242  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-760-475-2603

Query Match 95.1%; Score 39; DB 21; Length 242;  
Best Local Similarity 87.5%; Pred. No. 24;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTPLEELY 8  
|||||

Db 92 LTPLEELY 99

RESULT 5  
US-09-948-947-113  
; Sequence 113, Application US/09948947  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CL000786  
; CURRENT APPLICATION NUMBER: US/09/948,947  
; CURRENT FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/231,397  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 2172  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 113  
; LENGTH: 493  
; TYPE: PRT  
; ORGANISM: Human  
US-09-948-947-113

Query Match 95.1%; Score 39; DB 23; Length 493;  
Best Local Similarity 87.5%; Pred. No. 59;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTPLEELY 8  
|||||

Db 92 LTPLEELY 99

RESULT 6  
US-09-417-507-35531  
; Sequence 35531, Application US/09417507  
; GENERAL INFORMATION:  
; APPLICANT: KEITH G. WEINSTOCK ET AL.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS  
; TITLE OF INVENTION: FUMIGATUS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: PATH99-10  
; CURRENT APPLICATION NUMBER: US/09/417,507  
; CURRENT FILING DATE: 1999-10-14  
; NUMBER OF SEQ ID NOS: 44312

; SEQ ID NO 35531  
; LENGTH: 88  
; TYPE: PRT  
; ORGANISM: A.fumigatus  
US-09-417-507-35531

Query Match 90.2%; Score 37; DB 18; Length 88;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TPLEEY 8  
| | | | |  
Db 59 TPLEEY 65

RESULT 7  
US-60-316-362-3517  
; Sequence 3517, Application US/60316362  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M.  
; TITLE OF INVENTION: Identification of Essential Genes of Aspergillus fumigatus and Me  
; FILE REFERENCE: 10182-012-888  
; CURRENT APPLICATION NUMBER: US/60/316,362  
; CURRENT FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 4037  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3517  
; LENGTH: 310  
; TYPE: PRT  
; ORGANISM: Aspergillus fumigatus  
US-60-316-362-3517

Query Match 90.2%; Score 37; DB 24; Length 310;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TPLEEY 8  
| | | | |  
Db 123 TPLEEY 129

RESULT 8  
US-09-543-681A-4430  
; Sequence 4430, Application US/09543681A  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 4430  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-4430

Query Match 90.2%; Score 37; DB 19; Length 338;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TPLEEY 8  
| | | | |

Db 272 TPLEEY 278

RESULT 9  
PCT-US00-26524B-4656  
; Sequence 4656, Application PC/TUS0026524B  
; GENERAL INFORMATION:  
; APPLICANT: Birse et. al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept  
; FILE REFERENCE: PA005PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/26524B  
; CURRENT FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4656  
; LENGTH: 284  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US00-26524B-4656

Query Match 87.8%; Score 36; DB 1; Length 284;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTPLEEY 8  
: | | | | :  
Db 104 ITPLEDLY 111

RESULT 10  
US-09-328-352-7160  
; Sequence 7160, Application US/09328352  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTE  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7160  
; LENGTH: 806  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7160

Query Match 87.8%; Score 36; DB 17; Length 806;  
Best Local Similarity 75.0%; Pred. No. 4.5e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LTPLEEY 8  
: | | | | :  
Db 644 ITPLEEVY 651

RESULT 11  
US-60-146-055-775  
; Sequence 775, Application US/60146055  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID  
; FILE REFERENCE: CL000067  
; CURRENT APPLICATION NUMBER: US/60/146,055  
; CURRENT FILING DATE: 1999-07-28  
; NUMBER OF SEQ ID NOS: 938  
; SOFTWARE: FastSeq for Windows Version 3.0

```

; SEQ ID NO 775
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Drosophila
US-60-146-055-775

Query Match      85.4%; Score 35; DB 24; Length 71;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTPLEELY 8
    :|||||:
Db 46 ITPLEELF 53

RESULT 12
US-60-147-189-849
; Sequence 849, Application US/60147189
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; FILE REFERENCE: CLO00072
; CURRENT APPLICATION NUMBER: US/60/147,189
; CURRENT FILING DATE: 1999-08-08
; NUMBER OF SEQ ID NOS: 1115
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 849
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Drosophila
US-60-147-189-849

Query Match      85.4%; Score 35; DB 24; Length 71;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTPLEELY 8
    :|||||:
Db 46 ITPLEELF 53

RESULT 13
US-60-147-189-849
; Sequence 849, Application US/60147189
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; FILE REFERENCE: CLO00072
; CURRENT APPLICATION NUMBER: US/60/147,189
; CURRENT FILING DATE: 1999-08-08
; NUMBER OF SEQ ID NOS: 1115
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 849
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Drosophila
US-60-147-189-849

Query Match      85.4%; Score 35; DB 24; Length 71;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTPLEELY 8
    :|||||:
Db 46 ITPLEELF 53

RESULT 14
US-09-270-849B-189381
; Sequence 189381, Application US/09270849B
; GENERAL INFORMATION:
; APPLICANT: Swimmer et al.
; TITLE OF INVENTION: Insect genome survey devices
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/270,849B
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 195450
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 189381
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: at any location having Xaa, Xaa means any amino acid
US-09-270-849B-189381

Query Match      85.4%; Score 35; DB 16; Length 148;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPLEELY 8
    :||||:
Db 105 TPLEEYI 111

RESULT 15
US-09-270-767-44698
; Sequence 44698, Application US/09270767
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44698
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44698

Query Match      85.4%; Score 35; DB 16; Length 329;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPLEELY 8
    :||||:
Db 105 TPLEEYI 111

Search completed: January 29, 2002, 10:56:13
Job time: 2409 sec
```

```

; SEQ ID NO 775
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Drosophila
US-60-146-055-775

Query Match      85.4%; Score 35; DB 24; Length 71;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTPLEELY 8
    :|||||:
Db 46 ITPLEELF 53

RESULT 12
US-60-147-189-849
; Sequence 849, Application US/60147189
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; FILE REFERENCE: CLO00072
; CURRENT APPLICATION NUMBER: US/60/147,189
; CURRENT FILING DATE: 1999-08-08
; NUMBER OF SEQ ID NOS: 1115
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 849
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Drosophila
US-60-147-189-849

Query Match      85.4%; Score 35; DB 24; Length 71;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTPLEELY 8
    :|||||:
Db 46 ITPLEELF 53

RESULT 13
US-60-147-189-849
; Sequence 849, Application US/60147189
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; FILE REFERENCE: CLO00072
; CURRENT APPLICATION NUMBER: US/60/147,189
; CURRENT FILING DATE: 1999-08-08
; NUMBER OF SEQ ID NOS: 1115
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 849
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Drosophila
US-60-147-189-849

Query Match      85.4%; Score 35; DB 24; Length 71;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTPLEELY 8
    :|||||:
Db 46 ITPLEELF 53

RESULT 14
US-09-270-767-60152
; Sequence 60152, Application US/09270767
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60152
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-60152

Query Match      85.4%; Score 35; DB 16; Length 148;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPLEELY 8
    :||||:
Db 105 TPLEEYI 111
```



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:17 ; Search time 120.95 Seconds  
(without alignments)  
2.421 Million cell updates/sec

Title: US-09-763-397A-21  
Perfect score: 41  
Sequence: 1 LTPLEELY 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	78.0	89	5	US-09-620-394B-3470
2	32	78.0	106	5	US-09-708-427-40050
3	32	78.0	106	5	US-09-708-427-40054
4	32	78.0	117	5	US-09-620-394B-3469
5	32	78.0	122	5	US-09-620-394B-3468
6	32	78.0	134	5	US-09-708-427-40049
7	32	78.0	134	5	US-09-708-427-40053
8	32	78.0	139	5	US-09-708-427-40048
9	32	78.0	139	5	US-09-708-427-40052
10	32	78.0	435	5	US-09-801-368-150
11	32	78.0	435	5	US-09-487-558-150
12	31	75.6	224	6	US-10-015-127-11693
13	31	75.6	271	5	US-09-620-394B-3777
14	31	75.6	295	5	US-09-708-427-18806
15	31	75.6	2273	5	US-09-995-542-12
16	31	75.6	2310	5	US-09-995-542-10
17	30	73.2	153	5	US-09-708-427-26083
18	30	73.2	620	5	US-09-815-242-10088
19	30	73.2	620	5	US-09-815-242-14083
20	30	73.2	984	5	US-09-611-526-3757
21	30	73.2	1012	5	US-09-285-385C-4
22	30	73.2	1015	5	US-09-285-385C-2
23	30	73.2	1453	5	US-09-708-427-13383
24	30	73.2	1555	5	US-09-708-427-13382
25	30	73.2	1556	5	US-09-708-427-13381
26	29	70.7	186	5	US-09-910-150-14

27	29	70.7	237	5	US-09-916-790-17	Sequence 17, Appl
28	29	70.7	247	5	US-09-916-790-7	Sequence 7, Appl
29	29	70.7	260	7	US-60-335-352-4	Sequence 4, Appl
30	29	70.7	272	5	US-09-910-150-27	Sequence 27, Appl
31	29	70.7	273	5	US-09-922-138-11	Sequence 11, Appl
32	29	70.7	276	7	US-60-338-690-8	Sequence 8, Appl
33	29	70.7	278	5	US-09-842-582-4	Sequence 4, Appl
34	29	70.7	278	5	US-09-922-138-18	Sequence 18, Appl
35	29	70.7	278	5	US-09-922-138-27	Sequence 27, Appl
36	29	70.7	278	5	US-09-910-150-17	Sequence 17, Appl
37	29	70.7	278	5	US-09-910-150-31	Sequence 31, Appl
38	29	70.7	278	6	US-10-003-690-5	Sequence 5, Appl
39	29	70.7	278	7	US-60-338-690-9	Sequence 9, Appl
40	29	70.7	391	5	US-09-620-394B-7798	Sequence 7798, Ap
41	29	70.7	392	5	US-09-620-394B-7797	Sequence 7797, Ap
42	29	70.7	399	5	US-09-708-427-58997	Sequence 58997, A
43	29	70.7	404	5	US-09-620-394B-7796	Sequence 7796, Ap
44	29	70.7	468	5	US-09-708-427-58996	Sequence 58996, A
45	29	70.7	480	5	US-09-884-260A-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1  
US-09-620-394B-3470  
; Sequence 3470, Application US/09620394B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nickolai  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid  
; FILE REFERENCE: 2750-1067p  
; CURRENT APPLICATION NUMBER: US/09/620.394B  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 9131  
; SEQ ID NO 3470  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..89  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..89  
; OTHER INFORMATION: Ceres Seq. ID 1386864  
US-09-620-394B-3470

Query Match 78.0%; Score 32; DB 5; Length 89;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PLEELY 8  
Db 6 PLEELY 11

RESULT 2  
US-09-708-427-40050  
; Sequence 40050, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243p  
; CURRENT APPLICATION NUMBER: US/09/708.427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 40050  
; LENGTH: 106

```
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..106
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..106
; OTHER INFORMATION: Ceres Seq. ID 1842097
; OTHER INFORMATION: Ceres Seq. ID 1842097
US-09-708-427-40050

Query Match      78.0%; Score 32; DB 5; Length 106;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PLEELY 8
Db 6 PLEELY 11

RESULT 3
US-09-708-427-40054
; Sequence 40054, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40054
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..106
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..106
; OTHER INFORMATION: Ceres Seq. ID 1842101
US-09-708-427-40054

Query Match      78.0%; Score 32; DB 5; Length 106;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PLEELY 8
Db 6 PLEELY 11

RESULT 4
US-09-620-394B-3469
; Sequence 3469, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 3469
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: 1..117
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..117
; OTHER INFORMATION: Ceres Seq. ID 1386863
US-09-620-394B-3469

Query Match      78.0%; Score 32; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PLEELY 8
Db 34 PLEELY 39

RESULT 5
US-09-620-394B-3468
; Sequence 3468, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 3468
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..122
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..122
; OTHER INFORMATION: Ceres Seq. ID 1386862
US-09-620-394B-3468

Query Match      78.0%; Score 32; DB 5; Length 122;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PLEELY 8
Db 39 PLEELY 44

RESULT 6
US-09-708-427-40049
; Sequence 40049, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40049
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..134
; OTHER INFORMATION: Xaa is any amino acid
```



; NAME/KEY: misc\_feature  
; LOCATION: 1..134  
; OTHER INFORMATION: Ceres Seq. ID 1842096  
US-09-708-427-40049

Query Match 78.0%; Score 32; DB 5; Length 134;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PLEELY 8  
|||||  
Db 34 PLEELY 39

RESULT 7  
US-09-708-427-40053  
; Sequence 40053, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 40053  
; LENGTH: 134  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..134  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..134  
; OTHER INFORMATION: Ceres Seq. ID 1842100  
US-09-708-427-40053

Query Match 78.0%; Score 32; DB 5; Length 134;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PLEELY 8  
|||||  
Db 34 PLEELY 39

RESULT 8  
US-09-708-427-40048  
; Sequence 40048, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 40048  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..139  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..139  
; OTHER INFORMATION: Ceres Seq. ID 1842095

US-09-708-427-40048

Query Match 78.0%; Score 32; DB 5; Length 139;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PLEELY 8  
|||||  
Db 39 PLEELY 44

RESULT 9  
US-09-708-427-40052  
; Sequence 40052, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 40052  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..139  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..139  
; OTHER INFORMATION: Ceres Seq. ID 1842099  
US-09-708-427-40052

Query Match 78.0%; Score 32; DB 5; Length 139;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PLEELY 8  
|||||  
Db 39 PLEELY 44

RESULT 10  
US-09-801-368-150  
; Sequence 150, Application US/09801368  
; GENERAL INFORMATION:  
; APPLICANT: Busby, Robert  
; APPLICANT: Call, Brian  
; APPLICANT: Hecht, Peter  
; APPLICANT: Holtzman, Doug  
; APPLICANT: Madden, Kevin  
; APPLICANT: Maxon, Mary  
; APPLICANT: Milne, Todd  
; APPLICANT: Norman, Thea  
; APPLICANT: Royer, John  
; APPLICANT: Salama, Sofie  
; APPLICANT: Sherman, Amir  
; APPLICANT: Silva, Jeff  
; APPLICANT: Summers, Eric  
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
; FILE REFERENCE: 109272.147  
; CURRENT APPLICATION NUMBER: US/09/801,368  
; CURRENT FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 09/487,558  
; PRIOR FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US 60/160,587  
; PRIOR FILING DATE: 1999-10-20  
; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 150  
; LENGTH: 435  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-801-368-150

Query Match 78.0%; Score 32; DB 5; Length 435;  
Best Local Similarity 62.5%; Pred. No. 70;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTPLEELY 8  
|:||||:  
Db 90 LSPLEDIY 97

RESULT 11  
US-09-487-558-150  
; Sequence 150, Application US/09487558  
; GENERAL INFORMATION:  
; APPLICANT: Busby, Robert  
; APPLICANT: Cali, Brian  
; APPLICANT: Hecht, Peter  
; APPLICANT: Holtzman, Doug  
; APPLICANT: Madden, Kevin  
; APPLICANT: Maxon, Mary  
; APPLICANT: Milne, Todd  
; APPLICANT: Norman, Thea  
; APPLICANT: Royer, John  
; APPLICANT: Salama, Sofie  
; APPLICANT: Sherman, Amir  
; APPLICANT: Silva, Jeff  
; APPLICANT: Summers, Eric  
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production  
; FILE REFERENCE: 109272.147  
; CURRENT APPLICATION NUMBER: US/09/487,558  
; CURRENT FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US/09/801,368  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 09/487,558  
; PRIOR FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US 60/160,587  
; PRIOR FILING DATE: 1999-10-20  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 150  
; LENGTH: 435  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-487-558-150

Query Match 78.0%; Score 32; DB 5; Length 435;  
Best Local Similarity 62.5%; Pred. No. 70;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTPLEELY 8  
|:||||:  
Db 90 LSPLEDIY 97

RESULT 12  
US-10-015-127-11693  
; Sequence 11693, Application US/10015127  
; GENERAL INFORMATION:  
; APPLICANT: Bower, Stanley G.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Sphingomonas elodea genome sequences and uses thereof  
; FILE REFERENCE: 38-10(15806)B  
; CURRENT APPLICATION NUMBER: US/10/015,127

; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: US 60/252,455  
; PRIOR FILING DATE: 2000-11-22  
; NUMBER OF SEQ ID NOS: 14357  
; SEQ ID NO 11693  
; LENGTH: 224  
; TYPE: PRT  
; ORGANISM: Sphingomonas elodea  
US-10-015-127-11693

Query Match 75.6%; Score 31; DB 6; Length 224;  
Best Local Similarity 85.7%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TPLEELY 8  
|||  
Db 187 TPWEELY 193

RESULT 13  
US-09-620-394B-3777  
; Sequence 3777, Application US/09620394B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nickolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid  
; FILE REFERENCE: 2750-1067P  
; CURRENT APPLICATION NUMBER: US/09/620,394B  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 9131  
; SEQ ID NO 3777  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..271  
; OTHER INFORMATION: xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..271  
; OTHER INFORMATION: Ceres Seq. ID 1387554  
US-09-620-394B-3777

Query Match 75.6%; Score 31; DB 5; Length 271;  
Best Local Similarity 75.0%; Pred. No. 63;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LTPLEELY 8  
|:||||:  
Db 69 ITTLEELY 76

RESULT 14  
US-09-708-427-18806  
; Sequence 18806, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18806  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature

; LOCATION: 1..295  
; OTHER INFORMATION: xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..295  
; OTHER INFORMATION: Ceres Seq. ID 1834879  
US-09-708-427-18806

Query Match 75.6%; Score 31; DB 5; Length 295;  
Best Local Similarity 75.0%; Pred. No. 70;  
Matches 6; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

Qy 1 LTPLEELY 8  
Db 93 ITTLEELY 100

RESULT 15  
US-09-995-542-12  
; Sequence 12, Application US/09995542  
; GENERAL INFORMATION:  
; APPLICANT: Shutter, John  
; APPLICANT: Ullas, Laarni  
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and  
; FILE REFERENCE: 00-658-A  
; CURRENT APPLICATION NUMBER: US/09/995,542  
; CURRENT FILING DATE: 2001-11-28  
; PRIOR APPLICATION NUMBER: 60/253,520  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 2273  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-995-542-12

Query Match 75.6%; Score 31; DB 5; Length 2273;  
Best Local Similarity 71.4%; Pred. No. 8.3e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TPLEELY 8  
Db 1267 TPLEEIF 1273

Search completed: January 29, 2002, 10:58:18  
Job time: 2514 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:12 : Search time 1760.55 Seconds  
(without alignments)  
2.996 Million cell updates/sec

Title: US-09-763-397A-20

Perfect score: 96

Sequence: 1 NEREDERTLKKEYEDIVLK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pap:\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pap:\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pap:\*
- 4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pap:\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pap:\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pap:\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pap:\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pap:\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pap:\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pap:\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pap:\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pap:\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pap:\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pap:\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pap:\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pap:\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pap:\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pap:\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pap:\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pap:\*
- 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pap:\*
- 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pap:\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pap:\*
- 24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	19	6 US-08-295-164-5	Sequence 5, Appli
2	96	100.0	19	21 US-09-763-397A-20	Sequence 20, Appli
3	96	100.0	42	6 US-08-295-164-3	Sequence 3, Appli
4	96	100.0	43	3 US-07-606-259-5	Sequence 5, Appli
5	96	100.0	43	6 US-08-295-164-4	Sequence 4, Appli
6	96	100.0	350	21 US-09-763-397A-2	Sequence 2, Appli
7	96	100.0	1426	1 PCT-US94-10230-4	Sequence 4, Appli
8	96	100.0	1426	9 US-08-487-826-4	Sequence 4, Appli
9	96	100.0	1426	9 US-08-568-459-4	Sequence 4, Appli

10	96	100.0	1435	3 US-07-606-259-2	Sequence 2, Appli
11	96	100.0	1435	6 US-08-295-164-2	Sequence 2, Appli
12	96	100.0	1435	8 US-08-487-826A-4	Sequence 4, Appli
13	96	100.0	1435	16 US-09-210-288-4	Sequence 4, Appli
14	96	100.0	1604	10 US-08-605-185-3	Sequence 3, Appli
15	96	100.0	1786	7 US-08-307-742-3	Sequence 3, Appli
16	59	61.5	21	15 US-09-165-878-39	Sequence 39, Appli
17	54	56.2	21	15 US-09-165-878-72	Sequence 72, Appli
18	49	51.0	442	20 US-09-619-049-1140	Sequence 1140, Ap
19	49	51.0	442	24 US-60-167-217-22207	Sequence 22207, A
20	49	51.0	442	24 US-60-171-627-1764	Sequence 1764, Ap
21	49	51.0	442	24 US-60-173-464-18047	Sequence 18047, A
22	49	51.0	442	24 US-60-191-637-22043	Sequence 22043, A
23	49	51.0	442	24 US-60-191-681-17387	Sequence 17387, A
24	48	50.0	97	21 US-09-758-443-423	Sequence 423, App
25	48	50.0	236	1 PCT-US01-14827-8868	Sequence 8868, Ap
26	48	50.0	362	1 PCT-US99-22853B-173	Sequence 173, App
27	48	50.0	363	1 PCT-US99-22853B-172	Sequence 172, App
28	48	50.0	363	18 US-09-413-198-1783	Sequence 1783, A
29	48	50.0	549	24 US-60-324-109-21518	Sequence 21518, A
30	48	50.0	679	16 US-09-284-020-14	Sequence 14, Appl
31	48	50.0	685	16 US-09-284-020-16	Sequence 16, Appl
32	48	50.0	689	9 US-08-591-447-37	Sequence 37, Appl
33	48	50.0	689	9 US-08-591-447A-37	Sequence 37, Appl
34	48	50.0	689	16 US-09-284-020-2	Sequence 2, Appli
35	47	49.0	199	9 US-08-591-447-40	Sequence 40, Appl
36	47	49.0	199	9 US-08-591-447A-40	Sequence 40, Appl
37	47	49.0	200	9 US-08-591-447-17	Sequence 17, Appl
38	47	49.0	200	9 US-08-591-447A-17	Sequence 17, Appl
39	47	49.0	200	12 US-08-849-375-13	Sequence 13, Appl
40	47	49.0	465	24 US-60-312-544-8632	Sequence 8632, Ap
41	45	46.9	245	1 PCT-US01-00663-26532	Sequence 26532, A
42	45	46.9	245	22 US-09-864-761-33581	Sequence 33581, A
43	45	46.9	245	24 US-60-236-359-14953	Sequence 14953, A
44	45	46.9	259	17 US-09-328-352-7022	Sequence 7022, Ap
45	45	46.9	705	1 PCT-US01-14827-14972	Sequence 14972, A

ALIGNMENTS

RESULT 1  
US-08-295-164-5  
; Sequence 5, Application US/08295164  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Betty K  
; APPLICANT: Haynes, David J  
; APPLICANT: Orlandi, Palmar A  
; APPLICANT: Klotz, Francis W  
; APPLICANT: Carter, John M  
; APPLICANT: Camus, Daniel  
; APPLICANT: Zegans, Michael E  
; APPLICANT: Chuley, Jeffrey D  
; APPLICANT: Hadley, Terence J  
; TITLE OF INVENTION: Malaria Vaccine  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John Morank  
; STREET: HQ. USAMRDC, FORT DETRICK  
; CITY: Frederick  
; STATE: MD  
; COUNTRY: US  
; ZIP: 21702-5012  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/295,164  
; FILING DATE: 24-AUG-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:

Mon Feb 4 15:23:42 2002

```

; NAME: Hendricks, Glenna M
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Sims
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-7807
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; STRAIN: EBA-175
;
; US-08-295-164-5
;
Query Match 100.0%; Score 96; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEREDERTLKEYEDIVLK 19
DB 1 NEREDERTLKEYEDIVLK 19

RESULT 2
US-09-763-397A-20
; Sequence 20, Application US/09763397A
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis
; APPLICANT: Control and Prevention
; APPLICANT: Lal, Altaf A.
; APPLICANT: Ping Shu, Ya
; APPLICANT: Hasnain, Seyed E.
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci
; FILE REFERENCE: 6395-57049
; CURRENT APPLICATION NUMBER: US/09/763,397A
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/097,703
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: PCT / US99/18869
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; TYPE: PRT
; LENGTH: 19
; ORGANISM: Plasmodium falciparum
;
; US-09-763-397A-20
;
Query Match 100.0%; Score 96; DB 21; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEREDERTLKEYEDIVLK 19
DB 1 NEREDERTLKEYEDIVLK 19

RESULT 3
US-08-295-164-3
; Sequence 3, Application US/08295164
; GENERAL INFORMATION:
; APPLICANT: Sim, Betty K
; APPLICANT: Haynes, David J
;
; APPLICANT: Orlandi, Palmar A
; APPLICANT: Klotz, Francis W
; APPLICANT: Carter, John M
; APPLICANT: Camus, Daniel
; APPLICANT: Zegans, Michael E
; APPLICANT: Chuley, Jeffrey D
; APPLICANT: Hadley, Terence J
; TITLE OF INVENTION: Malaria Vaccine
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John Morank
; STREET: HQ. USAMRDC, FORT DETRICK
; CITY: Frederick
; STATE: MD
; COUNTRY: US
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/295,164
; FILING DATE: 24-AUG-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna M
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Sims
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-7807
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; STRAIN: EBA-175
;
; US-08-295-164-3
;
Query Match 100.0%; Score 96; DB 6; Length 42;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEREDERTLKEYEDIVLK 19
DB 8 NEREDERTLKEYEDIVLK 26

RESULT 4
US-07-606-259-5
; Sequence 5, Application US/07606259
; GENERAL INFORMATION:
; APPLICANT: SIM, Betty Kim Lee
; APPLICANT: ORLANDI, Palmar A.
; APPLICANT: HAYNES, JR., John David
; APPLICANT: KLOTZ, Francis W.
; APPLICANT: CARTER, John Mark
; APPLICANT: CAMUS, Daniel
; APPLICANT: ZEGHANS, Michael E.
; APPLICANT: CHULAY, Jeffrey E.
; APPLICANT: HADLEY, Terence J.
; TITLE OF INVENTION: Malaria Vaccine
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: U.S. Army Medical Research and Development  
ADDRESSEE: Command  
STREET: Fort Detrick  
CITY: Frederick  
STATE: Maryland  
COUNTRY: USA  
ZIP: 21702-5012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/606,259  
FILING DATE: 19901031  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/780,750  
FILING DATE: 27-SEP-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: MORAN, John F.  
REGISTRATION NUMBER: 26,313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301)619-2065  
TELEFAX: (301)619-2416  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
US-07-606-259-5

Query Match 100.0%; Score 96; DB 3; Length 43;  
Best Local Similarity 100.0%; Pred. No. 6.7e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NEREDERTLTKEYEDIVLK 19  
|||||  
Db 9 NEREDERTLTKEYEDIVLK 27

RESULT 5  
US-08-295-164-4  
; Sequence 4, Application US/08295164  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Betty K  
; APPLICANT: Haynes, David J  
; APPLICANT: Orlandi, Palmar A  
; APPLICANT: Klotz, Francis W  
; APPLICANT: Carter, John M  
; APPLICANT: Camus, Daniel  
; APPLICANT: Zegans, Michael E  
; APPLICANT: Chuley, Jeffrey D  
; APPLICANT: Hadley, Terence J  
; TITLE OF INVENTION: Malaria Vaccine  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John Morank  
; STREET: HQ. USAMRDC, FORT DETRICK  
; CITY: Frederick  
; STATE: MD  
; COUNTRY: US  
; ZIP: 21702-5012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/295,164  
FILING DATE: 24-AUG-1994  
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
NAME: Hendricks, Glenna M  
REGISTRATION NUMBER: 32,535  
REFERENCE/DOCKET NUMBER: Sims  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 619-7807  
TELEFAX: (301) 619-7714  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium falciparum  
STRAIN: EBA-175  
US-08-295-164-4

Query Match 100.0%; Score 96; DB 6; Length 43;  
Best Local Similarity 100.0%; Pred. No. 6.7e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NEREDERTLTKEYEDIVLK 19  
|||||  
Db 9 NEREDERTLTKEYEDIVLK 27

RESULT 6  
US-09-763-397A-2  
; Sequence 2, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Recombinant DNA/Protein  
US-09-763-397A-2

Query Match 100.0%; Score 96; DB 21; Length 350;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NEREDERTLTKEYEDIVLK 19  
|||||  
Db 223 NEREDERTLTKEYEDIVLK 241

RESULT 7  
PCT-US94-10230-4  
; Sequence 4, Application PC/TUS9410230

```
;
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 12
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; APPLICATION NUMBER: PCT/US94/10230
; FILING DATE: 07-SEP-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/119,677
; FILING DATE: 10-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1426 amino acids
; TYPE: amino acid
; STRANDEDNESS: Single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; PCT-US94-10230-4

Query Match 100.0%; Score 96; DB 1; Length 1426;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEREDERTLTKEYEDIVLK 19
Db 1066 NEREDERTLTKEYEDIVLK 1084

RESULT 8
US-08-487-826-4
; Sequence 4, Application US/08487826
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826
```

```
;
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1426 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-08-487-826-4

Query Match 100.0%; Score 96; DB 8; Length 1426;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEREDERTLTKEYEDIVLK 19
Db 1066 NEREDERTLTKEYEDIVLK 1084

RESULT 9
US-08-568-459-4
; Sequence 4, Application US/08568459
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,677
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 15280-139000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1426 amino acids
```



TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORGANISM: Plasmodium falciparum  
US-08-568-459-4

Query Match 100.0%; Score 96; DB 9; Length 1426;  
Best Local Similarity 100.0%; Pred. No. 6.2e-05;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEREDERTLTKEYEDIVLK 19  
|||||

Db 1066 NEREDERTLTKEYEDIVLK 1084

RESULT 10  
US-07-606-259-2  
; Sequence 2, Application US/07606259  
; GENERAL INFORMATION:  
; APPLICANT: SIM, Betty Kim Lee  
; APPLICANT: ORLANDI, Palmer A.  
; APPLICANT: HAYNES, JR., John David  
; APPLICANT: KLOTZ, Francis W.  
; APPLICANT: CARTER, John Mark  
; APPLICANT: CAMUS, Daniel  
; APPLICANT: ZEGHANS, Michael E.  
; APPLICANT: CHULAY, Jeffrey David  
; APPLICANT: HADLEY, Terence J.  
; TITLE OF INVENTION: Malaria Vaccine  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: U.S. Army Medical Research and Development  
; STREET: Command  
; CITY: Frederick  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 21702-5012  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07606,259  
; FILING DATE: 19901031  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/780,750  
; FILING DATE: 27-SEP-1985  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MORAN, John F.  
; REGISTRATION NUMBER: 26,313  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301)619-2065  
; TELEFAX: (301)619-2416  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1435 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-606-259-2

Query Match 100.0%; Score 96; DB 3; Length 1435;  
Best Local Similarity 100.0%; Pred. No. 6.3e-05;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEREDERTLTKEYEDIVLK 19  
|||||

Db 1069 NEREDERTLTKEYEDIVLK 1087

RESULT 11  
US-08-295-164-2  
; Sequence 2, Application US/08295164  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Betty K  
; APPLICANT: Haynes, David J  
; APPLICANT: Orlandi, Palmar A  
; APPLICANT: Klotz, Francis W  
; APPLICANT: Carter, John M  
; APPLICANT: Camus, Daniel  
; APPLICANT: Zegans, Michael E  
; APPLICANT: Chuley, Jeffrey D  
; APPLICANT: Hadley, Terence J  
; TITLE OF INVENTION: Malaria Vaccine  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John Morank  
; STREET: HQ. USAMRDC, FORT DETRICK  
; CITY: Frederick  
; STATE: MD  
; COUNTRY: US  
; ZIP: 21702-5012  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/295,164  
; FILING DATE: 24-AUG-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hendricks, Glenna M  
; REGISTRATION NUMBER: 32,535  
; REFERENCE/DOCKET NUMBER: Sims  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 619-7714  
; TELEFAX: (301) 619-7714  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1435 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Plasmodium falciparum  
; STRAIN: EBA-175  
US-08-295-164-2

Query Match 100.0%; Score 96; DB 6; Length 1435;  
Best Local Similarity 100.0%; Pred. No. 6.3e-05;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEREDERTLTKEYEDIVLK 19  
|||||

Db 1069 NEREDERTLTKEYEDIVLK 1087

RESULT 12  
US-08-487-826A-4  
; Sequence 4, Application US/08487826A  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.

```
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826A
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORGANISM: Plasmodium falciparum
;
; US-08-487-826A-4
;
; Query Match 100.0%; Score 96; DB 8; Length 1435;
; Best Local Similarity 100.0%; Pred. No. 6.3e-05;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 NEREDERTLTKYEDIVLK 19
; Db 1069 NEREDERTLTKYEDIVLK 1087
;
; RESULT 13
; US-09-210-288-4
; Sequence 4, Application US/09210288
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826A
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORGANISM: Plasmodium falciparum
;
; US-08-487-826A-4
;
; Query Match 100.0%; Score 96; DB 8; Length 1435;
; Best Local Similarity 100.0%; Pred. No. 6.3e-05;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 NEREDERTLTKYEDIVLK 19
; Db 1069 NEREDERTLTKYEDIVLK 1087
;
; RESULT 13
; US-09-210-288-4
; Sequence 4, Application US/09210288
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
;
; US-09-210-288-4
;
; Query Match 100.0%; Score 96; DB 16; Length 1435;
; Best Local Similarity 100.0%; Pred. No. 6.3e-05;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 NEREDERTLTKYEDIVLK 19
; Db 1069 NEREDERTLTKYEDIVLK 1087
;
; RESULT 14
; US-08-605-185-3
; Sequence 3, Application US/08605185
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION
; NUMBER OF SEQUENCES: 23
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,185
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/01900
; APPLICATION NUMBER: GB 93 18350.7
; FILING DATE: 03-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 94 17021.4
; FILING DATE: 23-AUG-1994
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1604 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
;
; US-08-605-185-3
;
; Query Match 100.0%; Score 96; DB 10; Length 1604;
; Best Local Similarity 100.0%; Pred. No. 7.2e-05;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 NEREDERTLTKEYEDIVLK 19  
 Db 1238 NEREDERTLTKEYEDIVLK 1256

RESULT 15

US-08-307-742-3  
 ; Sequence 3, Application US/08307742  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: ANTI-VIRAL FUSION PEPTIDES  
 ; NUMBER OF SEQUENCES: 6  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/307,742  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: WO PCT/GB93/00505  
 ; APPLICATION NUMBER: GB 9205276.0  
 ; FILING DATE: 11-MAR-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: GB 9214481.5  
 ; FILING DATE: 08-JUL-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: GB 9215829.4  
 ; FILING DATE: 24-JUL-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: GB 9219562.7  
 ; FILING DATE: 16-SEP-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: GB 9304311.5  
 ; FILING DATE: 03-MAR-1993  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1786 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: YES  
 US-08-307-742-3

Query Match 100.0%; Score 96; DB 7; Length 1786;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-05;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEREDERTLTKEYEDIVLK 19  
 Db 1420 NEREDERTLTKEYEDIVLK 1438

Search completed: January 29, 2002, 10:56:13  
 Job time: 2409 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:17 ; Search time 120.95 Seconds  
(without alignments)  
5.749 Million cell updates/sec

Title: US-09-763-397A-20  
Perfect score: 96  
Sequence: 1 NEREDERTLKREYDIVLK 19

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/plodata/2/paa/BCI\_NEW\_COMB.pep:\*  
2: /cgn2\_6/plodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/plodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/plodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/plodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/plodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/plodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	52.1	256	US-09-708-427-32947	Sequence 32947, A
2	50	52.1	362	US-09-708-427-32946	Sequence 32946, A
3	50	52.1	368	US-09-708-427-32945	Sequence 32945, A
4	46	47.9	376	US-09-708-427-24697	Sequence 24697, A
5	46	47.9	377	US-09-708-427-24696	Sequence 24696, A
6	46	46.9	384	US-09-708-427-24695	Sequence 24695, A
7	45	46.9	974	US-09-897-516-7722	Sequence 7722, Ap
8	42.5	44.3	1035	US-09-708-427-13441	Sequence 13441, A
9	42.5	44.3	1072	US-09-708-427-13440	Sequence 13440, A
10	42.5	44.3	1104	US-09-708-427-13439	Sequence 13439, A
11	41	42.7	198	US-09-897-516-7022	Sequence 13439, A
12	41	42.7	299	US-09-825-882-14	Sequence 14, Appl
13	41	42.7	871	US-10-000-823-7	Sequence 7, Appl1
14	41	42.7	1770	US-09-801-368-298	Sequence 298, Appl
15	41	42.7	1770	US-09-487-558-298	Sequence 298, Appl
16	40.5	42.2	297	US-09-708-427-32122	Sequence 32122, A
17	40.5	42.2	306	US-09-708-427-32121	Sequence 32121, A
18	40.5	42.2	399	US-09-708-427-32120	Sequence 32120, A
19	40	42.2	399	US-09-474-980B-156	Sequence 156, Appl
20	40	41.7	101	US-09-597-517-9	Sequence 9, Appl1
21	40	41.7	377	US-09-708-427-28019	Sequence 28019, A
22	40	41.7	388	US-09-897-516-8060	Sequence 8060, Ap
23	40	41.7	437	US-09-708-427-28018	Sequence 28018, A
24	40	41.7	437	US-09-935-144-40	Sequence 28018, A
25	40	41.7	478	US-09-708-427-28017	Sequence 28017, A
26	40	41.7	518	US-09-708-427-24763	Sequence 24763, A

27	40	41.7	534	US-09-708-427-24762	Sequence 24762, A
28	40	41.7	693	US-09-708-427-24761	Sequence 24761, A
29	40	41.7	1162	US-09-708-427-15454	Sequence 15454, A
30	40	41.7	1287	US-09-708-427-15453	Sequence 15453, A
31	40	41.7	1551	US-09-708-427-15452	Sequence 15452, A
32	39	40.6	101	US-09-605-703B-2194	Sequence 2194, Ap
33	39	40.6	120	US-09-620-111B-7367	Sequence 7367, Ap
34	39	40.6	142	US-09-620-111B-7370	Sequence 7370, Ap
35	39	40.6	200	US-09-921-496-84	Sequence 84, Appl
36	39	40.6	375	US-09-620-394B-6224	Sequence 6224, Ap
37	39	40.6	387	US-09-815-242-10293	Sequence 10293, A
38	39	40.6	415	US-09-708-427-902	Sequence 902, Appl
39	39	40.6	415	US-09-708-427-20902	Sequence 20902, A
40	39	40.6	415	US-09-708-427-76308	Sequence 76308, A
41	39	40.6	442	US-09-815-242-10904	Sequence 10904, A
42	39	40.6	450	US-09-708-427-901	Sequence 901, Appl
43	39	40.6	450	US-09-708-427-20901	Sequence 20901, A
44	39	40.6	450	US-09-708-427-76307	Sequence 76307, A
45	39	40.6	451	US-09-991-496-108	Sequence 108, Appl

## ALIGNMENTS

```

RESULT 1
US-09-708-427-32947
; Sequence 32947, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32947
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..256
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..256
; OTHER INFORMATION: Ceres Seq. ID 1835391
US-09-708-427-32947

```

Query Match 52.1% Score 50; DB 5; Length 256;  
Best Local Similarity 56.2% Pred. No. 2.8;  
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 NEREDERTLKREYDI 16
DB 58 NKKEDRSITREYKI 73

RESULT 2
US-09-708-427-32946
; Sequence 32946, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32946
; LENGTH: 362

```

```

; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..362
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..362
; OTHER INFORMATION: Ceres Seq. ID 1835390
US-09-708-427-32946
```

```

Query Match          52.1%; Score 50; DB 5; Length 362;
Best Local Similarity 56.2%; Pred. No. 4.1;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 NEREDERTLTKEYEDI 16
    |::||:|::||:|
Db 164 NRKEDORSLTRRETKI 179
```

```

RESULT 3
US-09-708-427-32945
; Sequence 32945, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32945
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..368
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..368
; OTHER INFORMATION: Ceres Seq. ID 1835389
US-09-708-427-32945
```

```

Query Match          52.1%; Score 50; DB 5; Length 368;
Best Local Similarity 56.2%; Pred. No. 4.2;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 NEREDERTLTKEYEDI 16
    |::||:|::||:|
Db 170 NRKEDORSLTRRETKI 185
```

```

RESULT 4
US-09-708-427-24697
; Sequence 24697, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24697
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
```

```

; NAME/KEY: misc_feature
; LOCATION: 1..376
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..376
; OTHER INFORMATION: Ceres Seq. ID 1812577
US-09-708-427-24697
```

```

Query Match          47.9%; Score 46; DB 5; Length 376;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 NEREDERTLTKEYE 14
    |::||:|::||:|
Db 164 NRKEDORSLTRRETE 177
```

```

RESULT 5
US-09-708-427-24696
; Sequence 24696, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24696
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..377
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..377
; OTHER INFORMATION: Ceres Seq. ID 1812576
US-09-708-427-24696
```

```

Query Match          47.9%; Score 46; DB 5; Length 377;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 NEREDERTLTKEYE 14
    |::||:|::||:|
Db 165 NRKEDORSLTRRETE 178
```

```

RESULT 6
US-09-708-427-24695
; Sequence 24695, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24695
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..384
; OTHER INFORMATION: Xaa is any amino acid
```

NAME/KEY: misc\_feature  
LOCATION: 1..384  
OTHER INFORMATION: Ceres Seq. ID 1812575  
US-09-708-427-24695

Query Match 47.9%; Score 46; DB 5; Length 384;  
Best Local Similarity 57.1%; Pred. No. 18;  
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 1 NREDERTLKEYE 14  
DB 172 NRKEDRSLTRTE 185

RESULT 7  
US-09-897-516-7722  
Sequence 7722, Application US/09897516  
GENERAL INFORMATION:  
APPLICANT: Corbin, David R.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Huesting, Joseph E.  
APPLICANT: Krasomil-Osterfeld, Karina C.  
APPLICANT: Slater, Steven C.  
APPLICANT: Spildonov, Sergei  
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51847)B  
CURRENT APPLICATION NUMBER: US/09/897,516  
CURRENT FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/215, 161  
PRIOR FILING DATE: 2000-06-30  
NUMBER OF SEQ ID NOS: 8409  
SEQ ID NO 7722  
LENGTH: 974  
TYPE: PRT  
ORGANISM: Xenorhabdus sp.  
US-09-897-516-7722

Query Match 46.9%; Score 45; DB 5; Length 974;  
Best Local Similarity 50.0%; Pred. No. 74;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
QY 2 EREDERTLKEYEDIV 17  
DB 904 ERKEDRLSEYQLI 919

RESULT 8  
US-09-708-427-13441  
Sequence 13441, Application US/09708427  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
FILE REFERENCE: 2750-1243P  
CURRENT APPLICATION NUMBER: US/09/708,427  
CURRENT FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 13441  
LENGTH: 1035  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
NAME/KEY: misc\_feature  
LOCATION: 1..1035  
OTHER INFORMATION: Xaa is any amino acid  
NAME/KEY: misc\_feature  
LOCATION: 1..1035  
OTHER INFORMATION: Ceres Seq. ID 1826362

US-09-708-427-13441

Query Match 44.3%; Score 42.5; DB 5; Length 1035;  
Best Local Similarity 62.5%; Pred. No. 1.9e+02;  
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;  
QY 2 ERE-DERTLKEYEDI 16  
DB 148 EREADEATLSKQREDL 163

RESULT 9  
US-09-708-427-13440  
Sequence 13440, Application US/09708427  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
FILE REFERENCE: 2750-1243P  
CURRENT APPLICATION NUMBER: US/09/708,427  
CURRENT FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 13440  
LENGTH: 1072  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
NAME/KEY: misc\_feature  
LOCATION: 1..1072  
OTHER INFORMATION: Xaa is any amino acid  
NAME/KEY: misc\_feature  
LOCATION: 1..1072  
OTHER INFORMATION: Ceres Seq. ID 1826361  
US-09-708-427-13440

Query Match 44.3%; Score 42.5; DB 5; Length 1072;  
Best Local Similarity 62.5%; Pred. No. 1.9e+02;  
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;  
QY 2 ERE-DERTLKEYEDI 16  
DB 185 EREADEATLSKQREDL 200

RESULT 10  
US-09-708-427-13439  
Sequence 13439, Application US/09708427  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
FILE REFERENCE: 2750-1243P  
CURRENT APPLICATION NUMBER: US/09/708,427  
CURRENT FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 13439  
LENGTH: 1104  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
NAME/KEY: misc\_feature  
LOCATION: 1..1104  
OTHER INFORMATION: Xaa is any amino acid  
NAME/KEY: misc\_feature  
LOCATION: 1..1104  
OTHER INFORMATION: Ceres Seq. ID 1826360  
US-09-708-427-13439

Query Match	44.3%	Score 42.5;	DB 5;	Length 1104;
Best Local Similarity	62.5%	Pred. No. 2e+02;		
Matches	10;	Conservative	3;	Mismatches
			2;	Indels
				1;
				Gaps
				1;
Qy	2	ERE-DEPTLTKYED	16	
Db	217	BRADDEATLSKORED	232	

```

RESULT 11
US-09-897-516-7022
: Sequence 7022, Application US/09897516
:
: GENERAL INFORMATION:
:
: APPLICANT: Corbin, David R.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Huesing, Joseph E.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Malvar, Thomas M.
: APPLICANT: Slater, Steven C.
: APPLICANT: Spiridonov, Sergei
:
: TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
:
: FILE REFERENCE: 38-21(51847)B
:
: CURRENT APPLICATION NUMBER: US/09/897,516
:
: CURRENT FILING DATE: 2001-06-29
:
: PRIOR APPLICATION NUMBER: US 60/215, 161
:
: PRIOR FILING DATE: 2000-06-30
:
: NUMBER OF SEQ ID NOS: 8409
:
: SEQ ID NO 7022
:
: LENGTH: 198
:
: TYPE: PRF
:
: ORGANISM: Xenorhabdus sp.
:
: US-09-897-516-7022

```

Query Match	42.7%	Score 41	DB 5	Length 198
Best Local Similarity	44.4%	Pred. No. 46		
Matches	8	Conservative	5	Mismatches 5
				Indels 0
				Gaps 0
QY	2	EREDERTLTKEEDLYLK	19	
	1	1181111111111111		
DB	148	ELWKRIKSTREYODLVTR	165	

```

RESULT 12 882-14
US-09-825-882-14
; Sequence 14, Application US/09825882
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON TILLOT
; TITLE OF INVENTION: T2R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003/0279152/RXT
; CURRENT APPLICATION NUMBER: US/09/825,882
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,532
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/247,014
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 299
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-825-882-14

```

Query Match	42.7%	Score 41:	DB 5:	length 299:
Best Local Similarity	80.0%	Pred. No. 74:		
Matches 8:	Conservative	0:	Mismatches 2:	Indels 0:
QY	5 DERTLTREYE 14			
Db	150 DERTVTRREYE 159			

```

RESULT 13
US-10-000-823-7
; Sequence 7, Application US/10000823
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN NUCLEIC ACID MOLECULES AND POLYPEPTIDES ENCODING A
; TITLE OF INVENTION: HUMAN ION CHANNEL EXPRESSED IN SPINAL CORD AND BRAIN
; FILE REFERENCE: DD10109NP
; CURRENT APPLICATION NUMBER: US/10/000,823
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,587
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 871
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-823-7

Query Match 42.7%; Score 41; DB 6; Length 871;
Best Local Similarity 42.1%; Pred. No. 2.6e+02;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 NERFDERTLKYEYEDIVLK 19
| || : || | : || |
Db 334 NTRENTKPYTKMYDILLK 352

```

```

RESULT 14
US-09-801-368-298
: Sequence 298, Application US/09801368
: GENERAL INFORMATION:
: APPLICANT: Busby, Robert
: APPLICANT: Cali, Brian
: APPLICANT: Hecht, Peter
: APPLICANT: Holtzman, Doug
: APPLICANT: Madden, Kevin
: APPLICANT: Maxon, Mary
: APPLICANT: Milne, Todd
: APPLICANT: Norman, Thea
: APPLICANT: Royer, John
: APPLICANT: Salama, Sofie
: APPLICANT: Sherman, Amir
: APPLICANT: Silva, Jeff
: APPLICANT: Summers, Eric
: TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
: FILE REFERENCE: 109272.147
: CURRENT APPLICATION NUMBER: US/09/801,368
: CURRENT FILING DATE: 2001-03-07
: PRIOR APPLICATION NUMBER: US 09/487,558
: PRIOR FILING DATE: 2000-01-19
: PRIOR APPLICATION NUMBER: US 60/160,587
: PRIOR FILING DATE: 1999-10-20
: NUMBER OF SEQ ID NOS: 440
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 298
: LENGTH: 1770
: TYPE: prt
: ORGANISM: Saccharomyces cerevisiae
US-09-801-368-298

```

Query Match	Similarity	42.7%	Score 41	DB 5	Length 1770
Best Local	Similarity	42.1%	Pred. No.	5.9e+02	
Matches	8	Conservative	4	Mismatches	7
				Indels	0
				Gaps	0
OY	1	NEREDERTLKKEYEDIVLK	19		
	11	11	11	11	11
Db	1212	NEEEREFLTPKADLIEK	1230		



RESULT 15

US-09-487-558-298

; Sequence 298, Application US/09487558  
; GENERAL INFORMATION:

; APPLICANT: Busby, Robert

; APPLICANT: Call, Brian

; APPLICANT: Hecht, Peter

; APPLICANT: Holtzman, Doug

; APPLICANT: Madden, Kevin

; APPLICANT: Maxon, Mary

; APPLICANT: Milne, Todd

; APPLICANT: Norman, Thea

; APPLICANT: Royer, John

; APPLICANT: Salama, Sofie

; APPLICANT: Sherman, Amir

; APPLICANT: Silva, Jeff

; APPLICANT: Summers, Eric

; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production

; TITLE OF INVENTION: In Fungi

; FILE REFERENCE: 109272.147

; CURRENT APPLICATION NUMBER: US/09/487,558

; CURRENT FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: US/09/801,368

; PRIOR FILING DATE: 2001-03-07

; PRIOR APPLICATION NUMBER: US 09/487,558

; PRIOR FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: US 60/160,587

; PRIOR FILING DATE: 1999-10-20

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 298

; LENGTH: 1770

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-487-558-298

Query Match 42.7%; Score 41; DB 5; Length 1770;

Best Local Similarity 42.1%; Pred. No. 5.9e+02;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 1 NEREDERTLTKEYEDIVLK 19

DB 1212 NEEEREPLTPRAKDLIEK 1230

Search completed: January 29, 2002, 10:58:17  
Job time: 2513 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:11 ; Search time 1760.55 seconds  
(without alignments)  
3.470 Million cell updates/sec

Title: US-09-763-397A-19  
Perfect score: 131  
Sequence: 1 EFTYMINFGQNYWEHPYQKS 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues  
Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main: \*  
1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	131	100.0	22	21	US-09-763-397A-19
2	131	100.0	350	21	US-09-763-397A-2
3	126	96.2	622	1	PCT-US92-02207-8
4	126	96.2	622	3	US-07-672-183A-8
5	126	96.2	622	4	US-08-072-867-8
6	126	96.2	622	4	US-08-075-783-9
7	57	43.5	89	24	US-60-207-216-871
8	57	43.5	99	1	PCT-US01-01307-606
9	57	43.5	103	1	PCT-US01-01307-859

10	57	43.5	374	1	PCT-US01-08631-48408	Sequence 48408, A
11	52	39.7	313	24	US-60-324-109-31002	Sequence 31002, A
12	52	39.7	506	20	US-09-672-785-2	Sequence 2, Appli
13	50	38.2	30	24	US-60-173-468-775	Sequence 775, App
14	50	38.2	30	24	US-60-173-468-775	Sequence 5435, Ap
15	49	37.4	512	20	US-09-672-785-4	Sequence 4, Appli
16	48	36.6	422	16	US-09-270-767-42572	Sequence 42572, A
17	48	36.6	503	24	US-60-173-464-19616	Sequence 19616, A
18	48	36.6	503	24	US-60-191-637-23927	Sequence 23927, A
19	48	36.6	503	24	US-60-191-681-18848	Sequence 18848, A
20	47	35.9	44	18	US-09-450-969-5226	Sequence 5226, Ap
21	47	35.9	103	22	US-09-834-366-24449	Sequence 24449, A
22	47	35.9	103	24	US-60-197-873-24449	Sequence 24449, A
23	47	35.9	258	16	US-09-248-796-25815	Sequence 25815, A
24	47	35.9	524	24	US-60-167-245-526	Sequence 526, App
25	47	35.9	524	24	US-60-173-464-11696	Sequence 11696, A
26	47	35.9	524	24	US-60-191-637-14390	Sequence 14390, A
27	47	35.9	524	24	US-60-191-681-11315	Sequence 11315, A
28	47	35.9	1024	19	US-09-562-737-87	Sequence 87, Appl
29	47	35.9	1620	24	US-60-252-625-2627	Sequence 2627, Ap
30	47	35.9	1620	24	US-60-257-931-3455	Sequence 3455, Ap
31	47	35.9	1620	24	US-60-259-308-4477	Sequence 4477, Ap
32	47	35.9	1627	16	US-09-232-991A-28697	Sequence 28697, A
33	46.5	35.5	151	24	US-60-173-464-24383	Sequence 24383, A
34	46.5	35.5	151	24	US-60-191-637-31024	Sequence 31024, A
35	46.5	35.5	151	24	US-60-191-681-24727	Sequence 24727, A
36	46	35.1	26	24	US-60-138-676-370	Sequence 370, App
37	46	35.1	26	24	US-60-141-862-719	Sequence 719, App
38	46	35.1	26	24	US-60-141-863-719	Sequence 719, App
39	46	35.1	26	24	US-60-143-993-1275	Sequence 1275, Ap
40	46	35.1	30	24	US-60-142-845-445	Sequence 445, App
41	46	35.1	151	16	US-09-212-152-1	Sequence 1, Appli
42	46	35.1	206	24	US-60-146-315-853	Sequence 853, App
43	46	35.1	265	21	US-09-760-469-1623	Sequence 1623, Ap
44	46	35.1	344	24	US-60-173-464-3743	Sequence 3743, Ap
45	46	35.1	344	24	US-60-191-637-4527	Sequence 4527, Ap

ALIGNMENTS

RESULT 1  
US-09-763-397A-19  
; Sequence 19, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 22  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-19

Query Match , 100.0%; Score 131; DB 21; Length 22;  
Best Local Similarity 100.0%; Pred. No. 6.4e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EFTYMINFGQNYWEHPYQKS 22

Db 1 EFTYMINFGRGQNYWHPYQKS 22  
|||||

RESULT 2  
US-09-763-397A-2

; Sequence 2, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Recombinant DNA/Protein  
US-09-763-397A-2

Query Match 100.0%; Score 131; DB 21; Length 350;  
Best Local Similarity 100.0%; Pred. No. 1.1e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWHPYQKS 22  
|||||

Db 242 EFTYMINFGRGQNYWHPYQKS 263

RESULT 3  
PCT-US92-02207-8

; Sequence 8, Application PC/TUS9202207  
; GENERAL INFORMATION:  
; APPLICANT: VIROGENETICS, CORPORATION  
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris, and Safford  
; ADDRESSEE: C/o William S. Frommer  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/02207  
; FILING DATE: 19-MAR-1992  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454310-2411  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712

; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 622 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
PCT-US92-02207-8

Query Match 96.2%; Score 126; DB 1; Length 622;  
Best Local Similarity 95.5%; Pred. No. 1.3e-10;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWHPYQKS 22  
|||||

Db 14 EFTYMINFGRGQNYWHPYQNS 35

RESULT 4

US-07-672-183A-8  
; Sequence 8, Application US/07672183A  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: de Taisne, Charles  
; APPLICANT: Tine, John A.  
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William S. Frommer  
; ADDRESSEE: C/o Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/672,183A  
; FILING DATE: 19910320  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454310-2300  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 622 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-07-672-183A-8

Query Match 96.2%; Score 126; DB 3; Length 622;  
Best Local Similarity 95.5%; Pred. No. 1.3e-10;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWHPYQKS 22  
|||||

Db 14 EFTYMINFGRGQNYWHPYQNS 35

RESULT 5

US-08-072-867-8  
; Sequence 8, Application US/08072867  
; GENERAL INFORMATION:

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-075-783-9

Query Match 96.2%; Score 126; DB 4; Length 622;
Best Local Similarity 95.5%; Pred. No. 1.3e-10;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGONYWEHPYQKS 22
| | | | | | | | | | | | | | | | | | | |
DB 14 EFTYMINFGRGONYWEHPYQNS 35

RESULT 7
US-60-207-216-871
; Sequence 871, Application US/60207216
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000580
; CURRENT APPLICATION NUMBER: US/60/207,216
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 871
; LENGTH: 89
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(89)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-207-216-871

Query Match 43.5%; Score 57; DB 24; Length 89;
Best Local Similarity 44.4%; Pred. No. 0.97;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 YMINFGRGONYWEHPYOK 21
| | | | | | | | | | | | | | | | | | | |
DB 65 YFENFANRQSIWDHPDK 82

RESULT 8
PCT-US01-01307-606
; Sequence 606, Application PC/TUS0101307
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT218PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01307
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1040
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 606
; LENGTH: 99

```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01307-606

Query Match      43.5%; Score 57; DB 1; Length 99;
Best Local Similarity 53.3%; Pred. No. 1.1;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY    4 YMINFGRGQNYWEHP 18
      |  ||  ||:  ||:  |
Db    54 YYFNFAQSQMWDPH 68

RESULT 9
PCT-US01-01307-859
; Sequence 859, Application PC/TUS0101307
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ18PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01307
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1040
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 859
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (70)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01307-859

Query Match      43.5%; Score 57; DB 1; Length 103;
Best Local Similarity 53.3%; Pred. No. 1.1;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY    4 YMINFGRGQNYWEHP 18
      |  ||  ||:  ||:  |
Db    54 YYFNFAQSQMWDPH 68

RESULT 10
PCT-US01-08631-48408
; Sequence 48408, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 48408
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (315)...(330)
; OTHER INFORMATION: WW DOMAIN SIGNATURE domain identified by eMATRIX, accession number PR00403B, p-value=3.898e-09, raw score of 12.19
```

```
; ORGANISM: Antirrhinum majus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of a protein having an activity to
; OTHER INFORMATION: directly convert flavanone to flavone
US-09-672-785-2

Query Match      39.7%; Score 52; DB 20; Length 506;
Best Local Similarity 42.1%; Pred. No. 36;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 FTYMINFGRGQNYWHPYQ 20
   | : || ||||| :
Db 391 FVNLMSGRPNYWNPM 409

RESULT 13
US-60-173-468-775
; Sequence 775, Application US/60173468
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS.
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO000181
; CURRENT APPLICATION NUMBER: US/60/173,468
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 1418
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 775
; LENGTH: 30
; TYPE: PRT
; ORGANISM: HUMAN
US-60-173-468-775

Query Match      38.2%; Score 50; DB 24; Length 30;
Best Local Similarity 46.7%; Pred. No. 3.9;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 YMINFGRGQNYWHP 18
   | || | : ||
Db 16 YFFNFANRQSIWDHP 30

RESULT 14
US-60-196-710-5435
; Sequence 5435, Application US/60196710
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS.
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00450
; CURRENT APPLICATION NUMBER: US/60/196,710
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7166
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5435
; LENGTH: 30
; TYPE: PRT
; ORGANISM: HUMAN
US-60-196-710-5435

Query Match      38.2%; Score 50; DB 24; Length 30;
Best Local Similarity 46.7%; Pred. No. 3.9;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 YMINFGRGQNYWHP 18
   | || | : ||
Db 16 YFFNFANRQSIWDHP 30
```

```
RESULT 15
US-09-672-785-4
; Sequence 4, Application US/09672785
; GENERAL INFORMATION:
; APPLICANT: Mizutani, Masako
; APPLICANT: Kasumi, Takaaki
; APPLICANT: Ayabe, Shin-ichi
; APPLICANT: Akashi, Tomoyoshi
; TITLE OF INVENTION: Genes Coding for Flavone Synthases
; FILE REFERENCE: 001560-383
; CURRENT APPLICATION NUMBER: US/09/672,785
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: PCT/JP00/04379
; PRIOR FILING DATE: 2000-01-30
; PRIOR APPLICATION NUMBER: PCT/JP00/00490
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: JP 11-205229
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: JP 11-22427
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Torenia hybrida
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of a protein having an activity to
; OTHER INFORMATION: directly convert flavanone to flavone
US-09-672-785-4
```

```
Query Match      37.4%; Score 49; DB 20; Length 512;
Best Local Similarity 42.1%; Pred. No. 1.1e+02;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 FTYMINFGRGQNYWHPYQ 20
   | : || ||||| :
Db 396 FVNLMSGRPNYWNPM 414
```

Search completed: January 29, 2002, 10:56:12  
Job time: 2408 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:16 ; Search time 120.95 seconds  
(without alignments)  
6.657 Million cell updates/sec

Title: US-09-763-397A-19  
Perfect score: 131  
Sequence: 1 EFTYINFGRGQNYWEHPYQKS 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New: \*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep: \*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep: \*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	35.9	292	5	US-09-708-427-15235
2	47	35.9	341	5	US-09-708-427-15234
3	47	35.9	452	5	US-09-708-427-15233
4	47	35.9	1620	5	US-09-815-242-5126
5	46	35.1	133	5	US-09-708-427-21796
6	46	35.1	230	5	US-09-708-427-21795
7	46	35.1	464	5	US-09-815-242-10647
8	46	35.1	467	5	US-09-815-242-4997
9	45	34.4	631	5	US-09-689-730-17
10	44	33.6	228	5	US-09-708-427-15512
11	43	32.8	219	5	US-09-800-198-91
12	43	32.8	266	5	US-09-800-198-90
13	43	32.8	266	5	US-09-876-225-3
14	43	32.8	415	5	US-09-800-198-89
15	42	32.1	135	5	US-09-708-427-67426
16	42	32.1	138	5	US-09-708-427-67425
17	42	32.1	572	5	US-09-708-427-9458
18	42	32.1	584	5	US-09-910-174A-16
19	42	32.1	615	5	US-09-708-427-9457
20	42	32.1	736	5	US-09-708-427-9456
21	42	32.1	2209	6	US-10-017-754-1903
22	41.5	31.7	582	1	PCT-US01-24104A-84
23	41.5	31.7	582	5	US-09-689-730-1
24	41	31.3	291	5	US-09-708-427-25013
25	41	31.3	372	7	US-60-325-537-15
26	41	31.3	665	5	US-09-708-427-15743

27	41	31.3	736	5	US-09-708-427-15742	Sequence 15742, A
28	41	31.3	821	5	US-09-708-427-15741	Sequence 15741, A
29	41	31.3	1143	5	US-09-708-427-10593	Sequence 10593, A
30	41	31.3	1180	5	US-09-708-427-10592	Sequence 10592, A
31	41	31.3	1254	5	US-09-708-427-10591	Sequence 10591, A
32	41	31.3	1977	5	US-09-976-594-757	Sequence 757, Appl
33	40.5	30.9	340	5	US-09-723-945-2	Sequence 2, Appl
34	40.5	30.9	489	5	US-09-689-730-11	Sequence 11, Appl
35	40.5	30.9	758	5	US-09-468-646A-31	Sequence 31, Appl
36	40.5	30.9	758	5	US-09-468-646B-31	Sequence 31, Appl
37	40	30.5	111	5	US-09-708-427-20343	Sequence 20343, A
38	40	30.5	136	5	US-09-708-427-23930	Sequence 23930, A
39	40	30.5	198	5	US-09-708-427-20342	Sequence 20342, A
40	40	30.5	300	5	US-09-953-499-10	Sequence 10, Appl
41	40	30.5	349	5	US-09-620-394B-1036	Sequence 1036, Ap
42	40	30.5	419	5	US-09-708-427-27392	Sequence 27392, A
43	40	30.5	435	5	US-09-708-427-27391	Sequence 27391, A
44	40	30.5	441	5	US-09-708-427-27390	Sequence 27390, A
45	40	30.5	453	5	US-09-897-516-7363	Sequence 7363, Ap

ALIGNMENTS

RESULT 1  
US-09-708-427-15235  
; Sequence 15235, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15235  
; LENGTH: 292  
; TYPE: PPT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..292  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..292  
; OTHER INFORMATION: Ceres Seq. ID 1828911  
US-09-708-427-15235

Query Match 35.9%; Score 47; DB 5; Length 292;  
Best Local Similarity 70.0%; Pred. No. 7.6;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 9 GRGQNYWEHP 18  
|||||:|:  
Db 11 GRGQNWDLDP 20

RESULT 2  
US-09-708-427-15234  
; Sequence 15234, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15234  
; LENGTH: 341

```

; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..341
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..341
; OTHER INFORMATION: Ceres Seq. ID 1828910
US-09-708-427-15234

Query Match          35.9%; Score 47; DB 5; Length 341;
Best Local Similarity 70.0%; Pred. No. 8.9;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 GRGQNYWHP 18
|||||:|
Db 60 GRGQNWDL 69

RESULT 3
US-09-708-427-15233
; Sequence 15233, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15233
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..452
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..452
; OTHER INFORMATION: Ceres Seq. ID 1828909
US-09-708-427-15233

Query Match          35.9%; Score 47; DB 5; Length 452;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 GRGQNYWHP 18
|||||:|
Db 171 GRGQNWDL 180

RESULT 4
US-09-815-242-5126
; Sequence 5126, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21

; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..133
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..133
; OTHER INFORMATION: Ceres Seq. ID 1840051
US-09-708-427-21796

Query Match          35.1%; Score 46; DB 5; Length 133;
Best Local Similarity 46.2%; Pred. No. 4.8;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 GRGQNYWHPYQK 21
| |: ||| :|
Db 13 GSGEAFWEHEWK 25

RESULT 6
US-09-708-427-21795
; Sequence 21795, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5126

Query Match          35.9%; Score 47; DB 5; Length 1620;
Best Local Similarity 36.8%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 EFTYMINFGRGQNYWHPY 19
||| : ||| :
Db 597 EFPYRLRHONGREYWIHDF 615

RESULT 5
US-09-708-427-21796
; Sequence 21796, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21796
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..133
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..133
; OTHER INFORMATION: Ceres Seq. ID 1840051
US-09-708-427-21796

Query Match          35.1%; Score 46; DB 5; Length 133;
Best Local Similarity 46.2%; Pred. No. 4.8;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 GRGQNYWHPYQK 21
| |: ||| :|
Db 13 GSGEAFWEHEWK 25

RESULT 6
US-09-708-427-21795
; Sequence 21795, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
```

; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21795  
; LENGTH: 230  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..230  
; OTHER INFORMATION: xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..230  
; OTHER INFORMATION: Ceres Seq. ID 1840050  
US-09-708-427-21795

Query Match 35.1%; Score 46; DB 5; Length 230;  
Best Local Similarity 46.2%; Pred. No. 8.4;  
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 GRGQNYWEHPYQK 21  
| : : : : :  
DB 110 GSGEAFWEHEWEK 122

RESULT 7  
US-09-815-242-10647  
; Sequence 10647, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Cart, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011a  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10647  
; LENGTH: 464  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-815-242-10647

Query Match 35.1%; Score 46; DB 5; Length 464;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 GRGQNYWEHPYQKS 22  
| : : : : :  
DB 28 GRGQNIWDFWYQEA 41

RESULT 8  
US-09-815-242-4997  
; Sequence 4997, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Cart, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011a  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4997  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-815-242-4997

Query Match 35.1%; Score 46; DB 5; Length 467;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 GRGQNYWEHPYQKS 22  
| : : : : :  
DB 31 GRGQNIWDFWYQEA 44

RESULT 9  
US-09-689-730-17  
; Sequence 17, Application US/09689730  
; GENERAL INFORMATION:  
; APPLICANT: SEIKI, Motoharu  
; APPLICANT: SATO, Hiroshi  
; APPLICANT: SHINAGAWA, Akira  
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR  
; FILE REFERENCE: 55-290P  
; CURRENT APPLICATION NUMBER: US/09/689,730  
; CURRENT FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: US/08/448,489  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 631  
; TYPE: PRT  
; ORGANISM: Unknown

```
;
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Known Member of
; OTHER INFORMATION: Matrix Metalloproteinase Family
US-09-689-730-17

Query Match      34.4%; Score 45; DB 5; Length 631;
Best Local Similarity 58.8%; Pred. No. 34;
Matches 10; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

QY 1 EFTYMINFGRCQNYWEH 17
   | | | | | | | | | |
Db 137 EADIMINEGR----WEH 149

RESULT 10
US-09-708-427-15512
; Sequence 15512, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15512
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..228
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..228
; OTHER INFORMATION: Ceres Seq. ID 1829311
US-09-708-427-15512

Query Match      33.6%; Score 44; DB 5; Length 228;
Best Local Similarity 54.5%; Pred. No. 17;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 GONYWEHPYOK 21
   | | | | | | | | | |
Db 110 GNFWEHEWEK 120

RESULT 11
US-09-800-198-91
; Sequence 91, Application US/09800198
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 91
; LENGTH: 219
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-800-198-91

Query Match      32.8%; Score 43; DB 5; Length 219;
Best Local Similarity 38.9%; Pred. No. 23;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 YMINFGRCQNYWEHPYOK 21
   | | | | | | | | | |
Db 161 FICNYGPGGNYPTWPKR 178

RESULT 12
US-09-800-198-90
; Sequence 90, Application US/09800198
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-198-90

Query Match      32.8%; Score 43; DB 5; Length 266;
Best Local Similarity 38.9%; Pred. No. 28;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 YMINFGRCQNYWEHPYOK 21
   | | | | | | | | | |
Db 171 FICNYGPGGNYPTWPKR 188

RESULT 13
US-09-876-225-3
; Sequence 3, Application US/09876225
; GENERAL INFORMATION:
; APPLICANT: Baylor College of Medicine
; APPLICANT: Thompson, Timothy C.
; APPLICANT: Ren, Chengzhen
; TITLE OF INVENTION: RTVP Based Compositions and Methods for the Treatment of Prostat
; FILE REFERENCE: 38594.0027
; CURRENT APPLICATION NUMBER: US/09/876,225
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/209,989
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-876-225-3

Query Match      32.8%; Score 43; DB 5; Length 266;
Best Local Similarity 38.9%; Pred. No. 28;
```

[illegible]



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:10 ; Search time 1760.55 seconds  
(without alignments)  
2.996 Million cell updates/sec

Title: US-09-763-397A-18  
Perfect score: 105  
Sequence: 1 DQPKQVEQHLTDYEIKIEG 19  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 3148936 seqs, 277657034 residues  
Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	105	100.0	19	21	US-09-763-397A-18
2	105	100.0	350	21	US-09-763-397A-2
3	105	100.0	622	1	PCT-US92-02207-8
4	105	100.0	622	3	US-07-672-183A-8
5	105	100.0	622	4	US-08-072-867-8
6	105	100.0	622	4	US-08-075-783-9
7	53	50.5	149	24	US-60-192-737-1135
8	53	50.5	149	24	US-60-194-106-704
9	53	50.5	160	24	US-60-192-737-1317

10	53	50.5	160	24	US-60-194-106-808
11	53	50.5	360	24	US-60-221-679-2
12	53	50.5	471	24	US-60-213-847-1176
13	53	50.5	568	1	PCT-US00-02544-4
14	53	50.5	568	1	PCT-US00-30542-8
15	53	50.5	568	17	US-09-371-674-4
16	53	50.5	568	21	US-09-706-426-8
17	53	50.5	568	22	US-09-855-145-4
18	53	50.5	588	1	PCT-US00-02544-2
19	53	50.5	588	1	PCT-US00-30542-6
20	53	50.5	588	15	US-09-152-814-1
21	53	50.5	588	17	US-09-371-674-2
22	53	50.5	588	21	US-09-706-426-6
23	53	50.5	588	22	US-09-855-145-2
24	51	48.6	226	13	US-08-984-099-3
25	50.5	48.1	681	18	US-09-494-810-2
26	50.5	48.1	681	18	US-09-494-810A-2
27	50.5	48.1	681	24	US-60-126-214-2
28	49	46.7	235	18	US-09-489-039A-9599
29	48	45.7	363	17	US-09-328-352-5693
30	46	43.8	77	24	US-60-215-161-6793
31	46	43.8	118	1	PCT-US01-08631-37013
32	46	43.8	278	19	US-09-573-655A-1732
33	45	42.9	165	22	US-09-898-554-24
34	45	42.9	207	22	US-09-898-554-26
35	45	42.9	227	24	US-60-324-109-25756
36	45	42.9	269	24	US-60-324-109-26164
37	45	42.9	302	24	US-60-312-544-5497
38	45	42.9	302	24	US-60-324-109-26768
39	45	42.9	363	22	US-09-898-554-20
40	44	41.9	66	24	US-60-173-466-55
41	44	41.9	95	1	PCT-US01-08631-34518
42	44	41.9	178	21	US-09-758-475-403
43	44	41.9	193	24	US-60-173-464-24239
44	44	41.9	194	24	US-60-191-637-30882
45	44	41.9	194	24	US-60-191-681-24571

ALIGNMENTS

RESULT 1  
US-09-763-397A-18  
; Sequence 18, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-18

Query Match 100.0% Score 105; DB 21; Length 19;  
Best Local Similarity 100.0%; Pred. No. 4.8e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DQPKQVEQHLTDYEIKIEG 19

Mon Feb 4 15:23:40 2002

us-09-763-397a-18.rapm

```
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US92-02207-8

Query Match      100.0%; Score 105; DB 1; Length 622;
Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKYEQHLTDYEKIREG 19
Db 348 DQPKYEQHLTDYEKIREG 366

RESULT 4
US-07-672-183A-8
; Sequence 8, Application US/07672183A
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: de Taisne, Charles
; APPLICANT: Tine, John A.
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William S. Frommer
; ADDRESSEE: c/o Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/672.183A
; FILING DATE: 19910320
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-672-183A-8

Query Match      100.0%; Score 105; DB 3; Length 622;
Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKYEQHLTDYEKIREG 19
Db 348 DQPKYEQHLTDYEKIREG 366

RESULT 5
US-08-072-867-8
; Sequence 8, Application US/08072867
; GENERAL INFORMATION:
```

```
Db 1 DQPKYEQHLTDYEKIREG 19

US-09-763-397A-2
; Sequence 2, Application US/09763397A
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis
; APPLICANT: Control and Prevention
; APPLICANT: Lal, Altaf A.
; APPLICANT: Ping Shi, Ya
; APPLICANT: Hasnain, Seyed E.
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci
; FILE REFERENCE: 6395-57049
; CURRENT APPLICATION NUMBER: US/09/763,397A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/097,703
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: PCT / US99/18869
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant DNA/Protein
US-09-763-397A-2

Query Match      100.0%; Score 105; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQPKYEQHLTDYEKIREG 19
Db 264 DQPKYEQHLTDYEKIREG 282

RESULT 3
PCT-US92-02207-8
; Sequence 8, Application PC/ATUS9202207
; GENERAL INFORMATION:
; APPLICANT: VIROGENETICS, CORPORATION
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris, and Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02207
; FILING DATE: 19-MAR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2411
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
```



APPLICANT: Paoletti, et al  
TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
NUMBER OF SEQUENCES: 74  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris, and Safford  
STREET: c/o William S. Frommer  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/072.867  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/852,305  
FILING DATE: 18-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2410  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 622 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-072-867-8

Query Match 100.0%; Score 105; DB 4; Length 622;  
Best Local Similarity 100.0%; Pred. No. 3.8e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQPKYEQHLDYKIKEG 19  
|||||  
Db 348 DQPKYEQHLDYKIKEG 366

RESULT 6  
US-08-075-783-9  
; Sequence 9, Application US/08075783  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: De Taisne, Charles  
; APPLICANT: Tine, John A.  
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; ADDRESSEE: c/o William S. Frommer  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/075.783  
; FILING DATE: 11-JUN-1993

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2500  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 622 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-075-783-9

Query Match 100.0%; Score 105; DB 4; Length 622;  
Best Local Similarity 100.0%; Pred. No. 3.8e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQPKYEQHLDYKIKEG 19  
|||||  
Db 348 DQPKYEQHLDYKIKEG 366

RESULT 7  
US-60-192-737-1135  
; Sequence 1135, Application US/60192737  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS OF THE  
; TITLE OF INVENTION: SER/THR AND TYR FAMILY OF KINASES, NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING THESE HUMAN KINASE PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CLO00402  
; CURRENT APPLICATION NUMBER: US/60/192.737  
; CURRENT FILING DATE: 2000-03-28  
; NUMBER OF SEQ ID NOS: 1342  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1135  
; LENGTH: 149  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-192-737-1135

Query Match 50.5%; Score 53; DB 24; Length 149;  
Best Local Similarity 64.3%; Pred. No. 6.7;  
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 KQYEQHLDYKIK 17  
|||  
Db 68 KQYKHLLTAYEKLE 81

RESULT 8  
US-60-194-106-704  
; Sequence 704, Application US/60194106  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CLO00413  
; CURRENT APPLICATION NUMBER: US/60/194.106  
; CURRENT FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 826  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 704  
; LENGTH: 149  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-194-106-704

Query Match 50.5%; Score 53; DB 24; Length 149;  
Best Local Similarity 64.3%; Pred. No. 6.7;  
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KQYEQHLTDYEKIK 17  
| | | : | | | : | | | :  
Db 68 KQYKHHLTAYEKL 81

RESULT 9  
US-60-192-737-1317  
; Sequence 1317, Application US/60192737  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS OF THE  
; TITLE OF INVENTION: SER/THR AND TYR FAMILY OF KINASES, NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING THESE HUMAN KINASE PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000402  
; CURRENT APPLICATION NUMBER: US/60/192,737  
; CURRENT FILING DATE: 2000-03-28  
; NUMBER OF SEQ ID NOS: 1342  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1317  
; LENGTH: 160  
; TYPE: PRT  
; ORGANISM: HUMAN  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(160)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-60-192-737-1317

Query Match 50.5%; Score 53; DB 24; Length 160;  
Best Local Similarity 64.3%; Pred. No. 7.3;  
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KQYEQHLTDYEKIK 17  
| | | : | | | : | | | :  
Db 85 KQYKHHLTAYEKL 98

RESULT 10  
US-60-194-106-808  
; Sequence 808, Application US/60194106  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL000413  
; CURRENT APPLICATION NUMBER: US/60/194,106  
; CURRENT FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 826  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 808  
; LENGTH: 160  
; TYPE: PRT  
; ORGANISM: HUMAN  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(160)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-60-194-106-808

Query Match 50.5%; Score 53; DB 24; Length 160;  
Best Local Similarity 64.3%; Pred. No. 7.3;  
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KQYEQHLTDYEKIK 17  
| | | : | | | : | | | :  
Db 85 KQYKHHLTAYEKL 98

Db 85 KQYKHHLTAYEKL 98  
| | | : | | | : | | | :  
RESULT 11  
US-60-221-679-2  
; Sequence 2, Application US/60221679  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Elliott, Vicki S.  
; APPLICANT: Ramkumar, Jaya  
; APPLICANT: Yao, Monique G.  
; APPLICANT: Burford, Neil  
; APPLICANT: Wang, Yu-mei E.  
; APPLICANT: Stewart, Elizabeth A.  
; APPLICANT: Gandhi, Ameena R.  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Lee, Ernestine A.  
; APPLICANT: Hafalia, April  
; APPLICANT: Lu, Dyung Aina M.  
; TITLE OF INVENTION: PROTEIN PHOSPHATASES  
; FILE REFERENCE: PI-0173 P  
; CURRENT APPLICATION NUMBER: US/60/221,679  
; CURRENT FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PERL Program  
; SEQ ID NO 2  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 7476908CD1  
US-60-221-679-2

Query Match 50.5%; Score 53; DB 24; Length 360;  
Best Local Similarity 62.5%; Pred. No. 20;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 PKQYEQHLTDYEKIKE 18  
| : | | | | | | | |  
Db 144 PEALKQHLQDYEKDKE 159

RESULT 12  
US-60-213-847-1176  
; Sequence 1176, Application US/60213847  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS OF THE  
; TITLE OF INVENTION: SER/THR AND TYR FAMILY OF KINASES, NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING THESE HUMAN KINASE PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000704  
; CURRENT APPLICATION NUMBER: US/60/213,847  
; CURRENT FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 1824  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1176  
; LENGTH: 471  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-213-847-1176

Query Match 50.5%; Score 53; DB 24; Length 471;  
Best Local Similarity 64.3%; Pred. No. 28;  
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KQYEQHLTDYEKIK 17  
| | | : | | | : | | | :  
Db 31 KQYKHHLTAYEKL 44





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:16 ; Search time 120.95 seconds  
(without alignments)  
5.749 Million cell updates/sec

Title: US-09-763-397A-18

Perfect score: 105

Sequence: 1 DQPKQYEQLHDYKIKEG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	53	50.5	551	6	US-10-016-333-24
2	46	43.8	77	5	US-09-897-516-6793
3	46	43.8	130	5	US-09-708-427-30615
4	46	43.8	222	5	US-09-708-427-30614
5	45	42.9	575	6	US-10-015-127-13974
6	42	40.0	587	5	US-09-708-427-33993
7	42	40.0	655	5	US-09-708-427-33992
8	42	40.0	715	7	US-60-325-537-18
9	41	39.0	256	5	US-09-897-516-8110
10	41	39.0	438	5	US-09-815-242-5240
11	41	39.0	445	5	US-09-815-242-12582
12	41	39.0	445	5	US-09-815-242-12940
13	41	39.0	540	5	US-09-611-526-2980
14	41	39.0	1485	5	US-09-897-516-7774
15	40	38.1	306	5	US-09-605-703B-2270
16	40	38.1	321	5	US-09-708-427-21965
17	40	38.1	355	5	US-09-708-427-21139
18	40	38.1	360	5	US-09-708-427-21964
19	40	38.1	412	6	US-10-015-127-12277
20	40	38.1	434	5	US-09-708-427-1306
21	40	38.1	486	5	US-09-708-427-1305
22	39.5	37.6	179	5	US-09-815-242-12256
23	39.5	37.6	388	5	US-09-815-242-5250
24	39.5	37.6	670	5	US-09-897-516-4390
25	39	37.1	75	5	US-09-708-427-15015
26	39	37.1	120	5	US-09-708-427-15014

27	39	37.1	182	5	US-09-561-263A-7	Sequence 7, Appl
28	39	37.1	182	5	US-09-561-263A-25	Sequence 25, Appl
29	39	37.1	220	5	US-09-708-427-50450	Sequence 50450, A
30	39	37.1	248	5	US-09-760-446A-1656	Sequence 1656, Ap
31	39	37.1	298	5	US-09-760-446A-1192	Sequence 1192, Ap
32	39	37.1	385	5	US-09-708-427-50449	Sequence 50449, A
33	39	37.1	486	5	US-09-801-368-154	Sequence 154, App
34	39	37.1	486	5	US-09-487-558-154	Sequence 154, App
35	39	37.1	575	5	US-09-897-516-7930	Sequence 7930, Ap
36	39	37.1	613	5	US-09-708-427-26645	Sequence 26645, A
37	39	37.1	705	5	US-09-815-242-10604	Sequence 10604, A
38	39	37.1	723	7	US-60-333-726-1343	Sequence 1343, Ap
39	39	37.1	805	5	US-09-708-427-26644	Sequence 26644, A
40	39	37.1	824	5	US-09-708-427-5544	Sequence 5544, Ap
41	39	37.1	832	5	US-09-708-427-4042	Sequence 4042, Ap
42	39	37.1	838	5	US-09-708-427-4041	Sequence 4041, Ap
43	39	37.1	847	5	US-09-708-427-26643	Sequence 26643, A
44	39	37.1	849	5	US-09-708-427-5543	Sequence 5543, Ap
45	39	37.1	889	5	US-09-708-427-5542	Sequence 5542, Ap

ALIGNMENTS

RESULT 1  
US-10-016-333-24  
; Sequence 24, Application US/10016333  
; GENERAL INFORMATION:  
; APPLICANT: Ian Edward James

; APPLICANT: Michael W. Lark  
; APPLICANT: Tania Tamson Testa  
; APPLICANT: John Anthony Feild  
; APPLICANT: Caretha Lee Creasy  
; APPLICANT: Brian Michael Burns  
; APPLICANT: Kimberly A. Brun  
; APPLICANT: Richard James Chenery  
; APPLICANT: Harma Ellens  
; APPLICANT: Lin Yue  
; APPLICANT: Erding Hu  
; APPLICANT: Zunxuan Chen  
; APPLICANT: Todd Fredrickson  
; APPLICANT: Randall Forrest Smith  
; APPLICANT: Xiaotong Li  
; APPLICANT: Yuan Zhu  
; APPLICANT: Peter R. Young  
; APPLICANT: Peter C. McDonnell  
; APPLICANT: Wendy S. Halsey  
; APPLICANT: Han N. Trinh  
; APPLICANT: Alexander H. Taylor  
; APPLICANT: Donald M. Wojchowski  
; APPLICANT: Stephen A. Hughes  
; APPLICANT: David J. Powell  
; APPLICANT: Gary Christie  
; APPLICANT: Michael N. Cook  
; APPLICANT: Kathleen T. Gallagher  
; APPLICANT: Kristine Kay Kikly  
; APPLICANT: David Michalovich  
; APPLICANT: Warren J. Leonard  
; APPLICANT: Katsutoshi Ozaki  
; TITLE OF INVENTION: MAMMALIAN POLYPEPTIDES AND  
; TITLE OF INVENTION: POLYNUCLEOTIDES  
; FILE REFERENCE: GP-70779B-C1  
; CURRENT APPLICATION NUMBER: US/10/016,333  
; CURRENT FILING DATE: 2001-12-06  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24  
; LENGTH: 551  
; TYPE: PRT  
; ORGANISM: MUS MUSCULUS

US-10-016-333-24

Query Match 50.5%; Score 53; DB 6; Length 551;  
Best Local Similarity 64.3%; Pred. No. 3.3;  
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KQYEQHLDYKIK 17  
| | | | | | | | | |  
Db 112 KQYKHLTAYEKL 125

RESULT 2  
US-09-897-516-6793  
; Sequence 6793, Application US/09897516  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Huesing, Joseph E.  
; APPLICANT: Krasomil-Osterfeld, Karina C.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)B  
; CURRENT APPLICATION NUMBER: US/09/897,516  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/215, 161  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8409  
; SEQ ID NO 6793  
; LENGTH: 77  
; TYPE: PRT  
; ORGANISM: Xenorhabdus sp.  
US-09-897-516-6793

Query Match 43.8%; Score 46; DB 5; Length 77;  
Best Local Similarity 50.8%; Pred. No. 3.8;  
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 YEQHLDYKIK 19  
| | | | | | | | | |  
Db 55 YKKHSTDFMKIRQ 68

RESULT 3  
US-09-708-427-30615  
; Sequence 30615, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30615  
; LENGTH: 130  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..130  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..130  
; OTHER INFORMATION: Ceres Seq. ID 1829600  
US-09-708-427-30615

Query Match 43.8%; Score 46; DB 5; Length 130;  
Best Local Similarity 61.5%; Pred. No. 6.9;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 YEQHLDYKIK 18  
| | | | | | | | | |  
Db 84 YAQHWTDQSK 96

RESULT 4  
US-09-708-427-30614  
; Sequence 30614, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30614  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..222  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..222  
; OTHER INFORMATION: Ceres Seq. ID 1829599  
US-09-708-427-30614

Query Match 43.8%; Score 46; DB 5; Length 222;  
Best Local Similarity 61.5%; Pred. No. 13;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 YEQHLDYKIK 18  
| | | | | | | | | |  
Db 176 YAQHWTDQSK 188

RESULT 5  
US-10-015-127-13974  
; Sequence 13974, Application US/10015127  
; GENERAL INFORMATION:  
; APPLICANT: Bower, Stanley G.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Spingomonas elodea genome sequences and uses thereof  
; FILE REFERENCE: 38-10(15806)B  
; CURRENT APPLICATION NUMBER: US/10/015,127  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: US 60/252,455  
; PRIOR FILING DATE: 2000-11-22  
; NUMBER OF SEQ ID NOS: 14357  
; SEQ ID NO 13974  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Spingomonas elodea  
US-10-015-127-13974

Query Match 42.9%; Score 45; DB 6; Length 575;  
Best Local Similarity 38.9%; Pred. No. 52;  
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 DQPKQYEQHLDYKIK 18  
| | | | | | | | | |  
Db 449 DLPENYEKHLAAFOFIRD 466

RESULT 6  
US-09-708-427-33993

; Sequence 33993, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 33993  
; LENGTH: 587  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..587  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..587  
; OTHER INFORMATION: Ceres Seq. ID 1837602  
US-09-708-427-33993

Query Match 40.0%; Score 42; DB 5; Length 587;  
Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 DDPKQYEQLHTDYEIKI 18  
| | | | | : | : | : |  
DB 36 DDPKQPSGPLTNYSKLVE 53

RESULT 7  
US-09-708-427-33992  
; Sequence 33992, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 33992  
; LENGTH: 655  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..655  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..655  
; OTHER INFORMATION: Ceres Seq. ID 1837601  
US-09-708-427-33992

Query Match 40.0%; Score 42; DB 5; Length 655;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 DDPKQYEQLHTDYEIKI 18  
| | | | | : | : | : |  
DB 104 DDPKQPSGPLTNYSKLVE 121

RESULT 8  
US-60-325-537-18  
; Sequence 18, Application US/60325537  
; GENERAL INFORMATION:  
; APPLICANT: Stein, Joshua C

; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Methods for Analyzing Biological Elements  
; FILE REFERENCE: 16517.246/38-10(52045)A  
; CURRENT APPLICATION NUMBER: US/60/325,537  
; CURRENT FILING DATE: 2001-10-01  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-60-325-537-18

Query Match 40.0%; Score 42; DB 7; Length 715;  
Best Local Similarity 46.7%; Pred. No. 1.9e+02;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 OYEQHLDYEDIKEG 19  
| | | | | : | : | : |  
DB 142 OYQHQVAMEAAKAG 156

RESULT 9  
US-09-897-516-8110  
; Sequence 8110, Application US/09897516  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Huesing, Joseph E.  
; APPLICANT: Krasomil-Osterfeld, Karina C.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)B  
; CURRENT APPLICATION NUMBER: US/09/897,516  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/215, 161  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8409  
; SEQ ID NO 8110  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Xenorhabdus sp.  
US-09-897-516-8110

Query Match 39.0%; Score 41; DB 5; Length 256;  
Best Local Similarity 53.8%; Pred. No. 81;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 KOYEQLHTDYEKI 16  
| | | | | : | : | : |  
DB 48 ROYDQMLADYNQI 60

RESULT 10  
US-09-815-242-5240  
; Sequence 5240, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haseibeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011a

```

; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5240
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-5240

Query Match          39.0%   Score 41; DB 5; Length 438;
Best Local Similarity 33.3%   Pred. No. 1.5e+02;
Matches 6; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY      1 DDPKQYEHLTDYEKIKE 18
Db      283 EQQRKHVEHLSSYYRFKQ 300
      : : : : : : : : : : : :
      : : : : : : : : : : : :

RESULT 11
US-09-815-242-12582
; Sequence 12582, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in

```

```

; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12582
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-12582

```



RESULT 15  
US-09-605-703B-2270  
; Sequence 2270, Application US/09605703B  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schroder, Hartwig

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:09 ; Search time 1760.55 Seconds  
(without alignments)  
2.839 Million cell updates/sec

Title: US-09-763-397A-17

Perfect score: 102

Sequence: 1 GNAEKYDKMDEPQHYGKS 18

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	102	100.0	18	21	US-09-763-397A-17	Sequence 17, Appl
2	102	100.0	350	21	US-09-763-397A-2	Sequence 2, Appli
3	93	91.2	622	1	PCT-US92-02207-8	Sequence 8, Appli
4	93	91.2	622	3	US-07-672-183A-8	Sequence 8, Appli
5	93	91.2	622	4	US-08-072-867-8	Sequence 8, Appli
6	93	91.2	622	4	US-08-075-783-9	Sequence 9, Appli
7	47	46.1	104	24	US-60-171-489-1218	Sequence 1218, Ap
8	44	43.1	393	18	US-09-417-507-36491	Sequence 36491, A
9	44	43.1	480	16	US-09-270-767-60056	Sequence 60056, A

10	44	43.1	480	16	US-09-270-849B-189090	Sequence 189090, A
11	44	43.1	634	16	US-09-252-991A-24400	Sequence 24400, A
12	44	43.1	635	16	US-09-270-767-44609	Sequence 44609, A
13	44	43.1	740	24	US-60-316-362-3482	Sequence 3482, Ap
14	43	42.2	133	24	US-60-215-161-6098	Sequence 6098, Ap
15	43	42.2	196	21	US-09-723-229-2	Sequence 2, Appli
16	43	42.2	226	24	US-60-191-637-27679	Sequence 27679, A
17	43	42.2	226	24	US-60-191-681-22288	Sequence 22288, A
18	43	42.2	327	24	US-60-167-245-604	Sequence 604, App
19	43	42.2	327	24	US-60-173-464-14129	Sequence 14129, A
20	43	42.2	339	24	US-60-191-637-17243	Sequence 17243, A
21	43	42.2	339	24	US-60-191-681-13652	Sequence 13652, A
22	43	42.2	508	1	PCT-US99-26796-44	Sequence 44, Appli
23	42.5	41.7	990	18	US-09-417-507-35717	Sequence 35717, A
24	42	41.2	405	24	US-60-128-476-4932	Sequence 4932, Ap
25	42	41.2	519	24	US-60-215-161-6859	Sequence 6859, Ap
26	42	41.2	554	19	US-09-540-236-2663	Sequence 2663, Ap
27	42	41.2	592	19	US-09-543-091-2	Sequence 2, Appli
28	42	41.2	622	17	US-09-328-352-7970	Sequence 7970, Ap
29	42	41.2	764	18	US-09-489-039A-8823	Sequence 8823, Ap
30	42	41.2	768	21	US-09-733-089-23126	Sequence 23126, A
31	42	41.2	768	22	US-09-816-660-23126	Sequence 23126, A
32	42	41.2	1443	21	US-09-733-089-19397	Sequence 19397, A
33	42	41.2	1443	22	US-09-816-660-19397	Sequence 19397, A
34	41.5	40.7	160	15	US-09-134-000-6652	Sequence 6652, Ap
35	41	40.2	187	1	PCT-US00-00724-710	Sequence 710, App
36	41	40.2	187	24	US-60-162-866-710	Sequence 15389, A
37	41	40.2	242	16	US-09-248-796-15389	Sequence 9416, Ap
38	41	40.2	284	24	US-60-312-544-9416	Sequence 1081, Ap
39	41	40.2	294	15	US-09-198-452A-1081	Sequence 1010, Ap
40	41	40.2	300	18	US-09-438-185-1010	Sequence 26726, A
41	41	40.2	363	16	US-09-252-991A-26726	Sequence 6123, Ap
42	41	40.2	948	17	US-09-328-352-6123	Sequence 6123, Ap
43	40.5	39.7	260	24	US-60-312-544-8243	Sequence 8243, Ap
44	40.5	39.7	271	24	US-60-312-544-8901	Sequence 8901, Ap
45	40.5	39.7	367	16	US-09-270-767-43442	Sequence 43442, A

#### ALIGNMENTS

RESULT 1

US-09-763-397A-17  
; Sequence 17, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-17

Query Match 100.0% Score 102: DB 21: Length 18;  
Best Local Similarity 100.0% Pred. No. 2.6e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GNAEKYDKMDEPQHYGKS 18

us-09-763-397a-17.rapm

Mon Feb 4 15:23:39 2002

```

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US92-02207-8

Query Match          91.2%; Score 93; DB 1; Length 622;
Best Local Similarity 94.4%; Pred. No. 5.7e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GNAEKYDKMDEPOHYGKS 18
Db 571 GNAEKYDKMDEPDYGKS 588

RESULT 4
US-07-672-183A-8
; Sequence 8, Application US/07672183A
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: de Taisne, Charles
; APPLICANT: Tine, John A.
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William S. Frommer
; ADDRESSEE: c/o Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/672,183A
; FILING DATE: 19910320
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-672-183A-8

Query Match          91.2%; Score 93; DB 3; Length 622;
Best Local Similarity 94.4%; Pred. No. 5.7e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GNAEKYDKMDEPOHYGKS 18
Db 571 GNAEKYDKMDEPDYGKS 588

RESULT 5
US-08-072-867-8
; Sequence 8, Application US/08072867
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis
; APPLICANT: Control and Prevention
; APPLICANT: Lal, Altaf A.
; APPLICANT: Ping Shi, Ya
; APPLICANT: Hasnain, Seyed E.
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci
; FILE REFERENCE: 6395-57049
; CURRENT APPLICATION NUMBER: US/09/763,397A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/097,703
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: PCT / US99/18869
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant DNA/Protein
US-09-763-397A-2

Query Match          100.0%; Score 102; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 9.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GNAEKYDKMDEPOHYGKS 18
Db 158 GNAEKYDKMDEPOHYGKS 175

RESULT 3
PCT-US92-02207-8
; Sequence 8, Application PC/TUS9202207
; GENERAL INFORMATION:
; APPLICANT: VIROGENETICS, CORPORATION
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris, and Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02207
; FILING DATE: 19-MAR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2411
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712

```

; APPLICANT: Paoletti, et al  
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris, and Safford  
; ADDRESSEE: c/o William S. Frommer  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/072,867  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/852,305  
; FILING DATE: 18-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454310-2410  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 622 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-072-867-8

Query Match 91.2%; Score 93; DB 4; Length 622;  
Best Local Similarity 94.4%; Pred. No. 5.7e-06;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GNAEKYDKMDEPQHYGKS 18  
|||||  
Db 571 GNAEKYDKMDEPQDYGKS 588

RESULT 6  
US-08-075-783-9  
; Sequence 9, Application US/08075783  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: De Taisne, Charles  
; APPLICANT: Tine, John A.  
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; ADDRESSEE: c/o William S. Frommer  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/075,783  
; FILING DATE: 11-JUN-1993

; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454310-2500  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 622 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-075-783-9

Query Match 91.2%; Score 93; DB 4; Length 622;  
Best Local Similarity 94.4%; Pred. No. 5.7e-06;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GNAEKYDKMDEPQHYGKS 18  
|||||  
Db 571 GNAEKYDKMDEPQDYGKS 588

RESULT 7  
US-60-171-489-1218  
; Sequence 1218, Application US/60171489  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL000165  
; CURRENT APPLICATION NUMBER: US/60/171,489  
; CURRENT FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 1412  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1218  
; LENGTH: 104  
; TYPE: PRT  
; ORGANISM: Human  
; US-60-171-489-1218

Query Match 46.1%; Score 47; DB 24; Length 104;  
Best Local Similarity 61.5%; Pred. No. 18;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 KYDKMDEPQHYGK 17  
|||  
Db 7 KWDLTPEQSYGK 19

RESULT 8  
US-09-417-507-36491  
; Sequence 36491, Application US/09417507  
; GENERAL INFORMATION:  
; APPLICANT: KEITH G. WEINSTOCK ET AL.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS  
; TITLE OF INVENTION: FUMIGATUS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: PATH99-10  
; CURRENT APPLICATION NUMBER: US/09/417,507  
; CURRENT FILING DATE: 1999-10-14  
; NUMBER OF SEQ ID NOS: 44312  
; SEQ ID NO 36491  
; LENGTH: 393  
; TYPE: PRT  
; ORGANISM: A.fumigatus  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (44).(53)

; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown  
US-09-417-507-36491

Query Match 43.1%; Score 44; DB 18; Length 393;  
Best Local Similarity 53.8%; Pred. No. 2.7e+02;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPQ 13  
|||||:::|  
Db 347 GNAKYAQVNPEN 359

RESULT 9  
US-09-270-767-60056  
; Sequence 60056, Application US/09270767  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 60056  
; LENGTH: 480  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-60056

Query Match 43.1%; Score 44; DB 16; Length 480;  
Best Local Similarity 60.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 YDKMDEPQHY 15  
||:::||||  
Db 360 YDRFERPQHY 369

RESULT 10  
US-09-270-849B-189090  
; Sequence 189090, Application US/09270849B  
; GENERAL INFORMATION:  
; APPLICANT: Swimmer et al.  
; TITLE OF INVENTION: Insect genome survey devices  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/270,849B  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 195450  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 189090  
; LENGTH: 480  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: at any location having Xaa, Xaa means any amino acid  
US-09-270-849B-189090

Query Match 43.1%; Score 44; DB 16; Length 480;  
Best Local Similarity 60.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 YDKMDEPQHY 15  
||:::||||  
Db 360 YDRFERPQHY 369

RESULT 11

US-09-252-991A-24400  
; Sequence 24400, Application US/09252991A  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24400  
; LENGTH: 634  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24400

Query Match 43.1%; Score 44; DB 16; Length 634;  
Best Local Similarity 70.0%; Pred. No. 4.9e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DKMDEPQHYG 16  
|:|||||  
Db 293 DGRDQPHYG 302

RESULT 12  
US-09-270-767-44609  
; Sequence 44609, Application US/09270767  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 44609  
; LENGTH: 635  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-44609

Query Match 43.1%; Score 44; DB 16; Length 635;  
Best Local Similarity 60.0%; Pred. No. 5e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 YDKMDEPQHY 15  
||:::||||  
Db 515 YDRFERPQHY 524

RESULT 13  
US-60-316-362-3482  
; Sequence 3482, Application US/60316362  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M.  
; TITLE OF INVENTION: Identification of Essential Genes of Aspergillus fumigatus and  
; FILE REFERENCE: 10182-012-888  
; CURRENT APPLICATION NUMBER: US/60/316,362  
; CURRENT FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 4037

; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3482  
; LENGTH: 740  
; TYPE: PRT  
; ORGANISM: Aspergillus fumigatus  
US-60-316-362-3482

Query Match 43.1%; Score 44; DB 24; Length 740;  
Best Local Similarity 53.8%; Pred. No. 6e+02;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEQ 13  
|||||:::|  
DB 548 GNAEKYAQVENPE 560

## RESULT 14

US-60-215-161-6098  
; Sequence 6098, Application US/60215161  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, David R.  
; APPLICANT: Corbin, Barry S.  
; APPLICANT: Huesing, Joseph E.  
; APPLICANT: Krasomil-Osterfeld, Karina C.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; APPLICANT: Hinkle, Gregory J.  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)A  
; CURRENT APPLICATION NUMBER: US/60/215,161  
; CURRENT FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8409  
; SEQ ID NO 6098  
; LENGTH: 133  
; TYPE: PRT  
; ORGANISM: Xenorhabdus sp.  
US-60-215-161-6098

Query Match 42.2%; Score 43; DB 24; Length 133;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 AEKDKMDEPQHYGKS 18  
|:::|::|:  
DB 61 AERHRASEEPHYNRS 76

## RESULT 15

US-09-723-229-2  
; Sequence 2, Application US/09723229  
; GENERAL INFORMATION:  
; APPLICANT: Douglas, James O.  
; TITLE OF INVENTION: Hypothalamic-Secreted Polypeptide  
; FILE REFERENCE: 01017/36937  
; CURRENT APPLICATION NUMBER: US/09/723,229  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: US 60/195,022  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 196  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-723-229-2

Query Match 42.2%; Score 43; DB 21; Length 196;  
Best Local Similarity 57.1%; Pred. No. 1.7e+02;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 EKYDKMDEPQHYGK 17  
|::|::|:  
DB 39 ENYDKYSEPRGYPK 52

Search completed: January 29, 2002, 10:56:10  
Job time: 2406 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:15 : Search time 120.95 Seconds  
(without alignments)  
5.446 Million cell updates/sec

Title: US-09-763-397A-17

Perfect score: 102

Sequence: 1 GNAEKYDKMDEPHYGS 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCr\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	43.1	572	5	US-09-648-004-8
2	43	42.2	133	5	US-09-897-516-6098
3	42	41.2	269	5	US-09-708-427-55948
4	42	41.2	281	5	US-09-708-427-55947
5	42	41.2	384	5	US-09-708-427-55946
6	42	41.2	519	5	US-09-897-516-6859
7	40.5	39.7	2011	5	US-09-832-292-31
8	40	39.2	697	5	US-09-815-242-12798
9	39	38.2	53	5	US-09-708-427-75548
10	39	38.2	139	5	US-09-708-427-74488
11	39	38.2	139	5	US-09-708-427-84690
12	39	38.2	158	5	US-09-708-427-73364
13	39	38.2	212	5	US-09-708-427-73363
14	39	38.2	246	5	US-09-708-427-73362
15	39	38.2	316	5	US-09-708-427-65583
16	39	38.2	330	5	US-09-708-427-65582
17	39	38.2	648	5	US-09-815-242-10504
18	39	38.2	1261	5	US-09-708-427-11963
19	38.5	37.7	954	5	US-09-708-427-17879
20	38.5	37.7	963	5	US-09-708-427-17878
21	38.5	37.7	964	5	US-09-708-427-17877
22	38	37.3	236	5	US-09-708-427-1259
23	38	37.3	267	5	US-09-708-427-1258
24	38	37.3	268	5	US-09-708-427-1257
25	38	37.3	533	6	US-10-015-127-10678
26	38	37.3	1144	5	US-09-708-427-15046

27	38	37.3	1304	5	US-09-708-427-15045	Sequence 15045, A
28	38	37.3	1313	5	US-09-708-427-15044	Sequence 15044, A
29	37	36.3	124	5	US-09-995-493-18	Sequence 18, Appl
30	37	36.3	232	5	US-09-708-427-33389	Sequence 33389, A
31	37	36.3	278	5	US-09-708-427-33388	Sequence 33388, A
32	37	36.3	308	5	US-09-708-427-33387	Sequence 33387, A
33	37	36.3	361	5	US-09-708-427-14814	Sequence 14814, A
34	37	36.3	384	5	US-09-708-427-14813	Sequence 14813, A
35	37	36.3	390	5	US-09-708-427-10424	Sequence 10424, A
36	37	36.3	412	5	US-09-708-427-14812	Sequence 14812, A
37	37	36.3	474	5	US-09-545-199D-81	Sequence 81, Appl
38	37	36.3	513	5	US-09-605-703B-1394	Sequence 1394, Ap
39	37	36.3	524	5	US-09-708-427-10423	Sequence 10423, A
40	37	36.3	544	5	US-09-759-272B-5	Sequence 5, Appl
41	37	36.3	587	5	US-09-708-427-10422	Sequence 10422, A
42	37	36.3	831	5	US-09-605-703B-1396	Sequence 1396, Ap
43	37	36.3	1003	5	US-09-989-723-33	Sequence 33, Appl
44	37	36.3	1003	5	US-09-989-724-33	Sequence 33, Appl
45	37	36.3	1003	5	US-09-989-730-33	Sequence 33, Appl

## ALIGNMENTS

```
RESULT 1
US-09-648-004-8
; Sequence 8, Application US/09648004
; GENERAL INFORMATION:
; APPLICANT: CHEN, QIONG
; APPLICANT: THOMAS, STUART
; APPLICANT: NAGARAAN, VASANTHA
; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND
; TITLE OF INVENTION: INTERMEDIATES
; FILE REFERENCE: CL-1341-A
; CURRENT APPLICATION NUMBER: US/09/648,004
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/252,553
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Acinetobacter sp.
US-09-648-004-8

Query Match 43.1% Score 44: DB 5: Length 572:
Best Local Similarity 64.3% Pred. No. 19:
Matches 9: Conservative 1: Mismatches 2: Indels 2: Gaps 1:

QY 5 KY--DKMDEPHYG 16
II II :|||||
DB 206 KYNLDKSDPHYG 219

RESULT 2
US-09-897-516-6098
; Sequence 6098, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesling, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spilidonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
```

;; PRIOR FILING DATE: 2000-06-30  
;; NUMBER OF SEQ ID NOS: 8409  
;; SEQ ID NO 6098  
;; LENGTH: 133  
;; TYPE: PRT  
;; ORGANISM: Xenorhabdus sp.  
US-09-897-516-6098

Query Match 42.2%; Score 43; DB 5; Length 133;  
Best Local Similarity 43.8%; Pred. No. 5.2;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 AERYDKMDEPOHYGK 18  
DB 61 AERRHASEEPTHYNRS 76

RESULT 3  
US-09-708-427-55948  
; Sequence 55948, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 55948  
; LENGTH: 269  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..269  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..269  
; OTHER INFORMATION: Ceres Seq. ID 1936382  
US-09-708-427-55948

Query Match 41.2%; Score 42; DB 5; Length 269;  
Best Local Similarity 41.2%; Pred. No. 17;  
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPOHYGK 17  
DB 148 GEGDKYDROAEKDHLSK 164

RESULT 4  
US-09-708-427-55947  
; Sequence 55947, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 55947  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..281  
; OTHER INFORMATION: Xaa is any amino acid

;; NAME/KEY: misc\_feature  
;; LOCATION: 1..281  
;; OTHER INFORMATION: Ceres Seq. ID 1936381  
US-09-708-427-55947

Query Match 41.2%; Score 42; DB 5; Length 281;  
Best Local Similarity 41.2%; Pred. No. 18;  
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPOHYGK 17  
DB 160 GEGDKYDROAEKDHLSK 176

RESULT 5  
US-09-708-427-55946  
; Sequence 55946, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 55946  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..384  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..384  
; OTHER INFORMATION: Ceres Seq. ID 1936380  
US-09-708-427-55946

Query Match 41.2%; Score 42; DB 5; Length 384;  
Best Local Similarity 41.2%; Pred. No. 26;  
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 GNAEKYDKMDEPOHYGK 17  
DB 263 GEGDKYDROAEKDHLSK 279

RESULT 6  
US-09-897-516-6859  
; Sequence 6859, Application US/09897516  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Huesting, Joseph E.  
; APPLICANT: Kitasomli-Osterfeld, Karina C.  
; APPLICANT: Matvar, Thomas M.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)B  
; CURRENT APPLICATION NUMBER: US/09/897,516  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/7215, 161  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8409  
; SEQ ID NO 6859  
; LENGTH: 519  
; TYPE: PRT  
; ORGANISM: Xenorhabdus sp.



```
; NAME/KEY: misc.feature
; LOCATION: 1..139
; OTHER INFORMATION: Ceres Seq. ID 1945476
US-09-708-427-74488
```

```
Query Match          38.2%; Score 39; DB 5; Length 139;
Best Local Similarity 53.8%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 2 NAEKYDKMDEPOH 14
:| | | | | | |
Db 114 HAEISYDWDIPDH 126
```

```
RESULT 11
US-09-708-427-84690
; Sequence 84690, Application US/09708427
```

```
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
```

```
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 84690
; LENGTH: 139
; TYPE: PRT
```

```
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
```

```
; LOCATION: 1..139
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..139
; OTHER INFORMATION: Ceres Seq. ID 1964740
US-09-708-427-84690
```

```
Query Match          38.2%; Score 39; DB 5; Length 139;
Best Local Similarity 53.8%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 2 NAEKYDKMDEPOH 14
:| | | | | | |
Db 114 HAEISYDWDIPDH 126
```

```
RESULT 12
US-09-708-427-73364
; Sequence 73364, Application US/09708427
```

```
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
```

```
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73364
; LENGTH: 158
; TYPE: PRT
```

```
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
```

```
; LOCATION: 1..158
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..158
; OTHER INFORMATION: Ceres Seq. ID 1943726
```

```
US-09-708-427-73364
```

```
Query Match          38.2%; Score 39; DB 5; Length 158;
Best Local Similarity 53.8%; Pred. No. 29;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 2 NAEKYDKMDEPOH 14
:| | | | | | |
Db 32 HAEISYDWDIPDH 44
```

```
RESULT 13
US-09-708-427-73363
; Sequence 73363, Application US/09708427
```

```
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
```

```
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
```

```
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73363
; LENGTH: 212
; TYPE: PRT
```

```
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
```

```
; LOCATION: 1..212
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..212
; OTHER INFORMATION: Ceres Seq. ID 1943725
US-09-708-427-73363
```

```
Query Match          38.2%; Score 39; DB 5; Length 212;
Best Local Similarity 53.8%; Pred. No. 41;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 2 NAEKYDKMDEPOH 14
:| | | | | | |
Db 86 HAEISYDWDIPDH 98
```

```
RESULT 14
US-09-708-427-73362
; Sequence 73362, Application US/09708427
```

```
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
```

```
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
```

```
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73362
; LENGTH: 246
; TYPE: PRT
```

```
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
```

```
; LOCATION: 1..246
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..246
; OTHER INFORMATION: Ceres Seq. ID 1943724
US-09-708-427-73362
```

Query Match 38.2%; Score 39; DB 5; Length 246;  
 Best Local Similarity 53.8%; Pred. No. 49;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY . 2 NAEKXDKMDEPOH 14  
 :|||:|  
 Db 120 HAEYDWMVDIPDH 132

RESULT 15

US-09-708-427-66583  
 ; Sequence 66583, Application US/09708427  
 ; GENERAL INFORMATION:  
 ; APPLICANT: N. ALEXANDROV et al.  
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
 ; TITLE OF INVENTION: THEREBY  
 ; FILE REFERENCE: 2750-1243P  
 ; CURRENT APPLICATION NUMBER: US/09/708,427  
 ; CURRENT FILING DATE: 2000-11-09  
 ; NUMBER OF SEQ ID NOS: 85364  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO: 66583  
 ; LENGTH: 316  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays subsp. mays  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: 1..316  
 ; OTHER INFORMATION: Xaa is any amino acid  
 ; NAME/KEY: misc.feature  
 ; LOCATION: 1..316  
 ; OTHER INFORMATION: Ceres Seq. ID 1930901  
 US-09-708-427-66583

Query Match 38.2%; Score 39; DB 5; Length 316;  
 Best Local Similarity 66.7%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 6 YDKMDEPOH 14  
 ||:|  
 Db 93 YDRYDEPXH 101

Search completed: January 29, 2002, 10:58:16  
 Job time: 2512 sec

**THIS PAGE BLANK (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: January 29, 2002, 10:56:08 ; Search time 1760.55 Seconds  
(without alignments)  
2.839 Million cell updates/sec

Title: US-09-763-397A-16

Perfect score: 100

Sequence: 1 DGNCEIPHVNEFSALDL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/US06\_COMB.pap.\*

2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pap.\*

3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pap.\*

4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pap.\*

5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pap.\*

6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pap.\*

7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pap.\*

8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pap.\*

9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pap.\*

10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pap.\*

11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pap.\*

12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pap.\*

13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pap.\*

14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pap.\*

15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pap.\*

16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pap.\*

17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pap.\*

18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pap.\*

19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pap.\*

20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pap.\*

21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pap.\*

22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pap.\*

23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pap.\*

24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	100	100.0	18	21	US-09-763-397A-16		Sequence 16, Appl
2	100	100.0	350	21	US-09-763-397A-2		Sequence 2, Appl
3	95	95.0	622	1	PCT-US92-02207-8		Sequence 8, Appl
4	95	95.0	622	3	US-07-672-183A-8		Sequence 8, Appl
5	95	95.0	622	4	US-08-072-867-8		Sequence 8, Appl
6	95	95.0	622	4	US-08-075-783-9		Sequence 9, Appl
7	51.5	51.5	301	16	US-09-248-796-18857		Sequence 18857, A
8	51.5	51.5	1042	24	US-60-314-050-7648		Sequence 7648, Ap
9	51	51.0	347	19	US-09-538-092-753		Sequence 753, App

10	47.5	47.5	160	20	US-09-689-980-1245	Sequence 1245, Ap
11	47.5	47.5	160	24	US-60-324-109-18127	Sequence 18127, A
12	47.5	47.5	177	24	US-60-324-109-25459	Sequence 25459, A
13	47.5	47.5	182	24	US-60-324-109-25168	Sequence 25168, A
14	47.5	47.5	192	20	US-09-689-980-1244	Sequence 1244, Ap
15	47	47.0	510	22	US-09-858-546-4	Sequence 4, Appl
16	46	46.0	194	24	US-60-175-871-1627	Sequence 1627, Ap
17	46	46.0	194	24	US-60-191-637-41784	Sequence 41784, A
18	46	46.0	194	24	US-60-191-700-1593	Sequence 1593, Ap
19	46	46.0	232	16	US-09-248-796-16149	Sequence 16149, A
20	46	46.0	352	24	US-60-259-128-5065	Sequence 5065, Ap
21	46	46.0	352	24	US-60-314-050-7189	Sequence 7189, Ap
22	46	46.0	371	24	US-60-212-664-586	Sequence 586, App
23	46	46.0	631	24	US-60-207-583-466	Sequence 466, App
24	45	45.0	56	1	PCT-US01-14827-8072	Sequence 8072, Ap
25	45	45.0	61	22	US-09-834-366-20159	Sequence 20159, A
26	45	45.0	61	24	US-60-197-873-20159	Sequence 20159, A
27	45	45.0	84	22	US-09-834-366-20619	Sequence 20619, A
28	45	45.0	84	24	US-60-197-873-20619	Sequence 20619, A
29	44	44.0	298	24	US-60-305-806-56	Sequence 56, Appl
30	44	44.0	1429	24	US-60-173-464-27182	Sequence 27182, A
31	44	44.0	1429	24	US-60-191-637-35218	Sequence 35218, A
32	44	44.0	1429	24	US-60-191-681-27637	Sequence 27637, A
33	44	44.0	1805	1	PCT-US98-00615-92	Sequence 92, Appl
34	44	44.0	1805	14	US-09-004-838-92	Sequence 92, Appl
35	43.5	43.5	115	20	US-09-643-672A-823	Sequence 823, App
36	43.5	43.5	118	24	US-60-324-109-20886	Sequence 20886, A
37	43.5	43.5	138	20	US-09-649-866A-482	Sequence 482, App
38	43.5	43.5	142	20	US-09-643-672A-822	Sequence 822, App
39	43.5	43.5	158	19	US-09-573-655A-390	Sequence 390, App
40	43.5	43.5	158	19	US-09-573-655A-1573	Sequence 1573, Ap
41	43.5	43.5	158	19	US-09-597-771-16	Sequence 16, Appl
42	43.5	43.5	159	19	US-09-570-581A-1835	Sequence 1835, Ap
43	43.5	43.5	160	19	US-09-597-771-12	Sequence 12, Appl
44	43.5	43.5	332	24	US-60-324-109-23277	Sequence 23277, A
45	43	43.0	85	18	US-09-417-507-26222	Sequence 26222, A

ALIGNMENTS

RESULT 1

US-09-763-397A-16

; Sequence 16, Application US/09763397A

; GENERAL INFORMATION:

; APPLICANT: The Government of the United States of America, as represented by the

; APPLICANT: Secretary of the Department of Health and Human Services, Centers for

; APPLICANT: Control and Prevention

; APPLICANT: Lal, Altaf A.

; APPLICANT: Ping Shi, Ya

; APPLICANT: Hasnain, Seved E.

; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa

; FILE REFERENCE: 6395-57049

; CURRENT APPLICATION NUMBER: US/09/763,397A

; CURRENT FILING DATE: 2001-02-16

; PRIOR FILING DATE: 1998-08-21

; PRIOR FILING DATE: 1998-08-21

; PRIOR FILING DATE: 1998-08-21

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 16

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Plasmodium falciparum

US-09-763-397A-16

Query Match 100.0%; Score 100; DB 21; Length 18;

Best Local Similarity 100.0%; Pred. No. 8.5e-10;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNCEIPHVNEFSALDL 18

Mon Feb 4 15:23:38 2002

```

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US92-02207-8

Query Match 95.0%; Score 95; DB 1; Length 622;
Best Local Similarity 94.4%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
Db 317 DGNCEIDPHVNEFPAIDL 334

RESULT 4
US-07-672-183A-8
; Sequence 8, Application US/07672183A
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: de Taisne, Charles
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William S. Frommer
; ADDRESS: c/o Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/672,183A
; FILING DATE: 19910320
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2300
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-672-183A-8

Query Match 95.0%; Score 95; DB 3; Length 622;
Best Local Similarity 94.4%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
Db 317 DGNCEIDPHVNEFPAIDL 334

RESULT 5
US-08-072-867-8
; Sequence 8, Application US/08072867
; GENERAL INFORMATION:

```

```

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US92-02207-8

Query Match 95.0%; Score 95; DB 1; Length 622;
Best Local Similarity 94.4%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
Db 317 DGNCEIDPHVNEFPAIDL 334

RESULT 4
US-07-672-183A-8
; Sequence 8, Application US/07672183A
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: de Taisne, Charles
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William S. Frommer
; ADDRESS: c/o Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/672,183A
; FILING DATE: 19910320
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2300
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-672-183A-8

Query Match 95.0%; Score 95; DB 3; Length 622;
Best Local Similarity 94.4%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DGNCEIDPHVNEFSAIDL 18
Db 317 DGNCEIDPHVNEFPAIDL 334

RESULT 5
US-08-072-867-8
; Sequence 8, Application US/08072867
; GENERAL INFORMATION:

```



; APPLICANT: Paoletti, et al  
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris, and Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/072,867  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/852,305  
; FILING DATE: 18-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454310-2410  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 622 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-072-867-8

Query Match 95.0%; Score 95; DB 4; Length 622;  
Best Local Similarity 94.4%; Pred. No. 4.3e-07;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNCEDPHVNEFSAIDL 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 317 DGNCEDPHVNEFPAIDL 334

RESULT 6  
US-08-075-783-9  
; Sequence 9, Application US/08075783  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: De Taisne, Charles  
; APPLICANT: Tine, John A.  
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/075,783  
; FILING DATE: 11-JUN-1993

; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454310-2500  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 622 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-075-783-9

Query Match 95.0%; Score 95; DB 4; Length 622;  
Best Local Similarity 94.4%; Pred. No. 4.3e-07;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNCEDPHVNEFSAIDL 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 317 DGNCEDPHVNEFPAIDL 334

RESULT 7  
US-09-248-796-18857  
; Sequence 18857, Application US/09248796  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI  
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796  
; CURRENT FILING DATE: 1999-02-12  
; NUMBER OF SEQ ID NOS: 28205  
; SEQ ID NO 18857  
; LENGTH: 301  
; TYPE: PRT  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (221),(244)  
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are un  
; US-09-248-796-18857

Query Match 51.5%; Score 51.5; DB 16; Length 301;  
Best Local Similarity 52.4%; Pred. No. 5.5;  
Matches 11; Conservative 2; Mismatches 3; Indels 5; Gaps 1;

QY 1 DGNCEDP-----HVNEFSAI 16  
| | | | | | | | | | | | | | | | | | | | | |  
Db 89 DFNCQDIPEFFEDHMLMAI 109

RESULT 8  
US-60-314-050-7648  
; Sequence 7648, Application US/60314050  
; GENERAL INFORMATION:  
; APPLICANT: Roemer, Terry  
; APPLICANT: Jiang, Bo  
; APPLICANT: Boone, Charles  
; APPLICANT: Bussey, Howard  
; APPLICANT: Ohlsen, Kari L.  
; TITLE OF INVENTION: GENE DISRUPTION METHODOLOGIES FOR DRUG TARGET DISCOVERY  
; FILE REFERENCE: 10182-013-888  
; CURRENT APPLICATION NUMBER: US/60/314,050  
; CURRENT FILING DATE: 2001-08-22  
; NUMBER OF SEQ ID NOS: 7834  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7648

```

; LENGTH: 1042
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (584)..(586)
; OTHER INFORMATION: X-any amino acid
; NAME/KEY: MISC_FEATURE
; LOCATION: (617)..(617)
; OTHER INFORMATION: X-any amino acid
; NAME/KEY: MISC_FEATURE
; LOCATION: (625)..(625)
; OTHER INFORMATION: X-any amino acid
US-60-314-050-7648

Query Match          51.5%; Score 51.5; DB 24; Length 1042;
Best Local Similarity 52.4%; Pred. No. 24;
Matches 11; Conservative 2; Mismatches 3; Indels 5; Gaps 1;

Qy 1 DGNCEIP-----HVNEFSAI 16
| | | | | | | | | | | | | |
Db 220 DENCQDIPEFFEDHNMELMAI 240

RESULT 9
US-09-538-092-753
; Sequence 753, Application US/09538092
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 753
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YOR262W
US-09-538-092-753

Query Match          51.0%; Score 51; DB 19; Length 347;
Best Local Similarity 61.5%; Pred. No. 8;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 6 DIPHVNEFSAILD 18
| : | | | | | | | |
Db 167 DLPHINVFSEKIDM 179

RESULT 10
US-09-689-980-1245
; Sequence 1245, Application US/09689980
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1237P
; CURRENT APPLICATION NUMBER: US/09/689,980
; CURRENT FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 3877
```

```

; SEQ ID NO 1245
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..160
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..160
; OTHER INFORMATION: Ceres Seq. ID 1597432
US-09-689-980-1245

Query Match          47.5%; Score 47.5; DB 20; Length 160;
Best Local Similarity 61.1%; Pred. No. 13;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

Qy 3 NCEDIPHYN--EFSAILD 18
| | | | | | | | | | | |
Db 81 NC-DIPHYNRTYQLIDI 97

RESULT 11
US-60-324-109-18127
; Sequence 18127, Application US/60324109
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)B
; CURRENT APPLICATION NUMBER: US/60/324,109
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 33196
; SEQ ID NO 18127
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
US-60-324-109-18127

Query Match          47.5%; Score 47.5; DB 24; Length 160;
Best Local Similarity 61.1%; Pred. No. 13;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

Qy 3 NCEDIPHYN--EFSAILD 18
| | | | | | | | | | | |
Db 81 NC-DIPHYNRTYQLIDI 97

RESULT 12
US-60-324-109-25459
; Sequence 25459, Application US/60324109
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)B
; CURRENT APPLICATION NUMBER: US/60/324,109
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 33196
; SEQ ID NO 25459
; LENGTH: 177
; TYPE: PRT
```

; ORGANISM: Zea mays  
; FEATURE:  
US-60-324-109-25459

Query Match 47.5%; Score 47.5; DB 24; Length 177;  
Best Local Similarity 61.1%; Pred. No. 14;  
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

QY 3 NCEDIPHVN--EFSIDL 18  
|| ||||| |:  
Db 98 NC-DIPHVNRTEYLIDI 114

RESULT 13  
US-60-324-109-25168  
; Sequence 25168, Application US/60324109  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Stein, Joshua  
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-10(52726)B  
; CURRENT APPLICATION NUMBER: US/60/324,109  
; CURRENT FILING DATE: 2001-09-21  
; NUMBER OF SEQ ID NOS: 33196  
; SEQ ID NO 25168  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
US-60-324-109-25168

Query Match 47.5%; Score 47.5; DB 24; Length 182;  
Best Local Similarity 61.1%; Pred. No. 15;  
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

QY 3 NCEDIPHVN--EFSIDL 18  
|| ||||| |:  
Db 103 NC-DIPHVNRTEYLIDI 119

RESULT 14  
US-09-689-980-1244  
; Sequence 1244, Application US/09689980  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1237P  
; CURRENT APPLICATION NUMBER: US/09/689,980  
; CURRENT FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 3877  
; SEQ ID NO 1244  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..192  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..192  
; OTHER INFORMATION: Ceres Seq. ID 1597431  
US-09-689-980-1244

Query Match 47.5%; Score 47.5; DB 20; Length 192;

Best Local Similarity 61.1%; Pred. No. 16;  
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

QY 3 NCEDIPHVN--EFSIDL 18  
|| ||||| |:  
Db 113 NC-DIPHVNRTEYLIDI 129

RESULT 15  
US-09-858-546-4  
; Sequence 4, Application US/09858546  
; GENERAL INFORMATION:  
; APPLICANT: SHAO, Wei et al  
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CL001232  
; CURRENT APPLICATION NUMBER: US/09/858,546  
; CURRENT FILING DATE: 2001-05-17  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 510  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-858-546-4

Query Match 47.0%; Score 47; DB 22; Length 510;  
Best Local Similarity 58.8%; Pred. No. 62;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 GNCEIPHVNEFSIDL 18  
||||| |:  
Db 335 GNCEFGHSNRCSYIDL 351

Search completed: January 29, 2002, 10:56:09  
Job time: 2405 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:15 ; Search time 120.95 Seconds  
(without alignments)  
5.446 Million cell updates/sec

Title: US-09-763-397A-16

Perfect score: 100

Sequence: 1 DGNCEIDPHVNEFSALDL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pcp.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pcp.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pcp.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pcp.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pcp.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pcp.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query % Match	Score	Length	DB ID	Description
1	47.5	47.5	135	5	US-09-708-427-49966
2	47.5	47.5	146	5	US-09-708-427-49942
3	47.5	47.5	149	5	US-09-708-427-79333
4	47.5	47.5	160	5	US-09-708-427-49467
5	47.5	47.5	160	7	US-60-337-358-461
6	47.5	47.5	163	5	US-09-708-427-49964
7	47.5	47.5	173	5	US-09-708-427-49941
8	45	45.0	283	5	US-09-605-703B-2242
9	43.5	43.5	158	5	US-09-708-427-15548
10	43	43.0	135	5	US-09-637-780B-576
11	43	43.0	138	5	US-09-637-780B-575
12	43	43.0	182	5	US-09-637-780B-574
13	43	43.0	679	5	US-09-708-427-32251
14	43	43.0	741	5	US-09-708-427-32250
15	43	43.0	827	5	US-09-708-427-32249
16	42	42.0	1579	5	US-09-801-368-368
17	42	42.0	1579	5	US-09-487-558-388
18	41.5	41.5	162	5	US-09-620-111B-2714
19	41.5	41.5	531	5	US-09-708-427-20569
20	41.5	41.5	549	5	US-09-708-427-20568
21	41.5	41.5	566	5	US-09-708-427-20567
22	41	41.0	660	5	US-09-976-594-787
23	41	41.0	1420	5	US-09-708-427-11605
24	41	41.0	1517	5	US-09-708-427-11604
25	41	41.0	1519	5	US-09-708-427-11603
26	41	41.0	1593	5	US-09-708-427-3888

ALIGNMENTS

RESULT 1

US-09-708-427-49966

; Sequence 49966, Application US/09708427

; GENERAL INFORMATION:

; APPLICANT: N. ALEXANDROV et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID

; FILE REFERENCE: 2750-1243P

; CURRENT APPLICATION NUMBER: US/09/708,427

; CURRENT FILING DATE: 2000-11-09

; NUMBER OF SEQ ID NOS: 85364

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 49966

; LENGTH: 135

; TYPE: PRT

; ORGANISM: Zea mays subsp. mays

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Xaa is any amino acid

; LOCATION: 1..135

; NAME/KEY: misc.feature

; LOCATION: 1..135

; OTHER INFORMATION: Ceres Seq. ID 1923498

US-09-708-427-49966

Query Match 47.5%; Score 47.5; DB 5; Length 135;  
Best Local Similarity 61.1%; Pred. No. 0.66; 2: Indels 3: Gaps 2:  
Matches 11; Conservative 2; Mismatches

QY 3 NCEDIPHN--EFSALDL 18

|| ||||| I: ||:

Db 81 NC-DIPHNRTXQLIDI 97

RESULT 2

US-09-708-427-49942

; Sequence 49942, Application US/09708427

; GENERAL INFORMATION:

; APPLICANT: N. ALEXANDROV et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID

; FILE REFERENCE: 2750-1243P

; CURRENT APPLICATION NUMBER: US/09/708,427

; CURRENT FILING DATE: 2000-11-09

; NUMBER OF SEQ ID NOS: 85364

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 49942

; LENGTH: 146

```
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..146
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..146
; OTHER INFORMATION: Ceres Seq. ID 1923441
US-09-708-427-49942
```

Query Match 47.5% Score 47.5; DB 5; Length 146;

Best Local Similarity 61.1%; Pred. No. 0.72; 2; Indels 3; Gaps 2;

```
Matches 11; Conservative 2; Mismatches 2;
Oy 3 NCEDIPHVN--EFSAILD 18
Db 81 NC-DIPHVNRTEYQLIDI 97
|| ||||| |:
```

## RESULT 3

```
US-09-708-427-79333
; Sequence 79333, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 79333
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..149
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..149
; OTHER INFORMATION: Ceres Seq. ID 1965146
US-09-708-427-79333
```

Query Match 47.5% Score 47.5; DB 5; Length 149;

Best Local Similarity 61.1%; Pred. No. 0.73; 2; Indels 3; Gaps 2;

```
Matches 11; Conservative 2; Mismatches 2;
Oy 3 NCEDIPHVN--EFSAILD 18
Db 40 NC-DIPHVNRTEYQLIDI 56
|| ||||| |:
```

## RESULT 4

```
US-09-708-427-49467
; Sequence 49467, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49467
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: 1..160
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..160
; OTHER INFORMATION: Ceres Seq. ID 1921936
US-09-708-427-49467
```

Query Match 47.5% Score 47.5; DB 5; Length 160;

Best Local Similarity 61.1%; Pred. No. 0.8; 2; Indels 3; Gaps 2;

```
Matches 11; Conservative 2; Mismatches 2;
Oy 3 NCEDIPHVN--EFSAILD 18
Db 81 NC-DIPHVNRTEYQLIDI 97
|| ||||| |:
```

## RESULT 5

```
US-60-337-358-461
; Sequence 461, Application US/60337358
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Chomet, Paul S.
; APPLICANT: Daly, Mackenzie
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Mollan
; APPLICANT: Dong, Jinzhuo
; APPLICANT: Duff, Stephen M.
; APPLICANT: Edgerton, Michael D.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshieh
; APPLICANT: Johnson, Richard G.
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A.
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmathi, Machikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Sheridan, Paul L.
; APPLICANT: Sherman, Paul L.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyuo
; APPLICANT: Xin, Zhanqun
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)A
; CURRENT APPLICATION NUMBER: US/60/337,358
; CURRENT FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 745
```

; SEQ ID NO 461  
; LENGTH: 160  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-60-337-358-461

Query Match 47.5%; Score 47.5; DB 7; Length 160;  
Best Local Similarity 61.1%; Pred. No. 0.8;  
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

QY 3 NCEDIPHVN--EFSIDL 18  
|| ||||| | : ||:  
Db 81 NC-DIPHVNRTEYQLIDI 97

## RESULT 6

US-09-708-427-49964  
; Sequence 49964, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 49964  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..163  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..163  
; OTHER INFORMATION: Ceres Seq. ID 1923496  
US-09-708-427-49964

Query Match 47.5%; Score 47.5; DB 5; Length 163;  
Best Local Similarity 61.1%; Pred. No. 0.81;  
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

QY 3 NCEDIPHVN--EFSIDL 18  
|| ||||| | : ||:  
Db 109 NC-DIPHVNRTEYQLIDI 125

## RESULT 7

US-09-708-427-49941  
; Sequence 49941, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 49941  
; LENGTH: 173  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..173  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..173

; OTHER INFORMATION: Ceres Seq. ID 1923440  
US-09-708-427-49941

Query Match 47.5%; Score 47.5; DB 5; Length 173;  
Best Local Similarity 61.1%; Pred. No. 0.87;  
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 2;

QY 3 NCEDIPHVN--EFSIDL 18  
|| ||||| | : ||:  
Db 108 NC-DIPHVNRTEYQLIDI 124

## RESULT 8

US-09-605-703B-2242  
; Sequence 2242, Application US/09605703B  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schroder, Hartwig  
; APPLICANT: Zelder, Oskar  
; APPLICANT: Haberhauer, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL  
; TITLE OF INVENTION: PROTEINS  
; FILE REFERENCE: BGI-129CP  
; CURRENT APPLICATION NUMBER: US/09/605,703B  
; CURRENT FILING DATE: 2000-06-27  
; PRIOR APPLICATION NUMBER: 60/142,764  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: 60/152,318  
; PRIOR FILING DATE: 1999-09-03  
; NUMBER OF SEQ ID NOS: 2934  
; SEQ ID NO 2242  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-605-703B-2242

Query Match 45.0%; Score 45; DB 5; Length 283;  
Best Local Similarity 63.6%; Pred. No. 4;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 DIPHVNEFSAI 16  
| ||||| |:  
Db 9 DAPHINEFPAL 19

## RESULT 9

US-09-708-427-15548  
; Sequence 15548, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15548  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..158  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..158  
; OTHER INFORMATION: Ceres Seq. ID 1829388  
US-09-708-427-15548

Query Match 43.5%; Score 43.5; DB 5; Length 158;  
Best Local Similarity 50.0%; Pred. No. 3.8;  
Matches 9; Conservative 4; Mismatches 2; Indels 3; Gaps 2;

QY 3 NCEDIPHN--EFSIDL 18  
|||:||||:|:  
DB 79 NC-DVPHVNRDYLQIDI 95

RESULT 10  
US-09-637-780B-576  
; Sequence 576, Application US/09637780B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1096P  
; CURRENT APPLICATION NUMBER: US/09/637,780B  
; CURRENT FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 1401  
; SEQ ID NO 576  
; LENGTH: 135  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: misc\_feature  
; LOCATION: 1..135  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..135  
; OTHER INFORMATION: Ceres Seq. ID 1482447  
US-09-637-780B-576

Query Match 43.0%; Score 43; DB 5; Length 135;  
Best Local Similarity 61.5%; Pred. No. 3.8;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 DIPHVNEFSIDL 18  
:::|||||:|:  
DB 2 ELPHVNLISKIDL 14

RESULT 11  
US-09-637-780B-575  
; Sequence 575, Application US/09637780B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1096P  
; CURRENT APPLICATION NUMBER: US/09/637,780B  
; CURRENT FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 1401  
; SEQ ID NO 575  
; LENGTH: 138  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: misc\_feature  
; LOCATION: 1..138  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..138  
; OTHER INFORMATION: Ceres Seq. ID 1482446  
US-09-637-780B-575

Query Match 43.0%; Score 43; DB 5; Length 138;  
Best Local Similarity 61.5%; Pred. No. 3.9;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 DIPHVNEFSIDL 18  
:::|||||:|:  
DB 5 ELPHVNLISKIDL 17

RESULT 12  
US-09-637-780B-574  
; Sequence 574, Application US/09637780B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1096P  
; CURRENT APPLICATION NUMBER: US/09/637,780B  
; CURRENT FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 1401  
; SEQ ID NO 574  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: misc\_feature  
; LOCATION: 1..182  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..182  
; OTHER INFORMATION: Ceres Seq. ID 1482445  
US-09-637-780B-574

Query Match 43.0%; Score 43; DB 5; Length 182;  
Best Local Similarity 61.5%; Pred. No. 5.4;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 DIPHVNEFSIDL 18  
:::|||||:|:  
DB 49 ELPHVNLISKIDL 61

RESULT 13  
US-09-708-427-32251  
; Sequence 32251, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32251  
; LENGTH: 679  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: misc\_feature  
; LOCATION: 1..679  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..679  
; OTHER INFORMATION: Ceres Seq. ID 1833855  
US-09-708-427-32251

Query Match 43.0%; Score 43; DB 5; Length 679;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CEDIPHV 10



Search completed: January 29, 2002, 10:58:15  
Job time: 2511 sec

```
Db      12 CEDIPHV 18
|||||||
RESULT 14
US-09-708-427-32250
; Sequence 32250, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32250
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..741
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..741
; OTHER INFORMATION: Ceres Seq. ID 1833854
US-09-708-427-32250

Query Match      43.0%; Score 43; DB 5; Length 741;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 CEDIPHV 10
|||||||
Db      74 CEDIPHV 80

RESULT 15
US-09-708-427-32249
; Sequence 32249, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32249
; LENGTH: 827
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..827
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..827
; OTHER INFORMATION: Ceres Seq. ID 1833853
US-09-708-427-32249

Query Match      43.0%; Score 43; DB 5; Length 827;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 CEDIPHV 10
|||||||
Db     160 CEDIPHV 166
```

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: January 29, 2002, 10:56:07 ; Search time 1760.55 seconds  
(without alignments)  
1.262 Million cell updates/sec

Title: US-09-763-397A-15  
Perfect score: 52  
Sequence: 1 GQGHMHG 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main.\*  
1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US00\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	8	21	US-09-763-397A-15
2	52	100.0	272	1	PCT-US99-26796-59
3	52	100.0	350	21	US-09-763-397A-2
4	43	82.7	198	23	US-09-902-540-15333
5	41	78.8	157	23	US-09-902-540-16719
6	41	78.8	269	1	PCT-US01-08631-41146
7	41	78.8	289	24	US-60-161-932-1547
8	41	78.8	363	24	US-60-167-217-241
9	41	78.8	363	24	US-60-173-464-228

10	41	78.8	449	24	US-60-167-324-1944	Sequence 1944, Ap
11	41	78.8	449	24	US-60-191-637-258	Sequence 258, App
12	41	78.8	449	24	US-60-191-637-26251	Sequence 26251, A
13	41	78.8	449	24	US-60-191-637-42246	Sequence 42246, A
14	41	78.8	449	24	US-60-191-681-213	Sequence 213, App
15	41	78.8	449	24	US-60-191-681-20860	Sequence 20860, A
16	41	78.8	518	24	US-60-173-464-19849	Sequence 19849, A
17	41	78.8	1209	24	US-60-173-464-19253	Sequence 19253, A
18	41	78.8	1331	24	US-60-161-932-2507	Sequence 2507, Ap
19	41	78.8	1419	24	US-60-191-637-23438	Sequence 23438, A
20	41	78.8	1419	24	US-60-191-681-18494	Sequence 18494, A
21	40	76.9	118	24	US-60-173-464-30035	Sequence 30035, A
22	40	76.9	119	24	US-60-191-637-39511	Sequence 39511, A
23	40	76.9	119	24	US-60-191-681-30544	Sequence 30544, A
24	40	76.9	211	16	US-09-252-991A-22035	Sequence 22035, A
25	40	76.9	358	18	US-09-417-507-30578	Sequence 30578, A
26	40	76.9	472	24	US-60-167-217-7880	Sequence 7880, Ap
27	40	76.9	472	24	US-60-173-464-6289	Sequence 6289, Ap
28	40	76.9	472	24	US-60-191-637-7785	Sequence 7785, Ap
29	40	76.9	472	24	US-60-191-681-6063	Sequence 6063, Ap
30	40	76.9	609	24	US-60-191-637-41277	Sequence 41277, A
31	40	76.9	609	24	US-60-191-700-1149	Sequence 1149, Ap
32	40	76.9	662	24	US-60-161-932-2573	Sequence 2573, Ap
33	39	75.0	102	7	US-08-390-207-36	Sequence 36, Appl
34	39	75.0	123	1	PCT-US00-26524B-6816	Sequence 6816, Ap
35	39	75.0	226	18	US-09-489-039A-10824	Sequence 10824, A
36	39	75.0	258	21	US-09-758-442-475	Sequence 475, App
37	39	75.0	280	17	US-09-328-352-7042	Sequence 7042, Ap
38	39	75.0	311	24	US-60-150-584-1015	Sequence 1015, Ap
39	39	75.0	366	24	US-60-167-217-14635	Sequence 14635, A
40	39	75.0	366	24	US-60-173-464-11873	Sequence 11873, A
41	39	75.0	366	24	US-60-191-637-14585	Sequence 14585, A
42	39	75.0	366	24	US-60-191-681-11480	Sequence 11480, A
43	39	75.0	376	1	PCT-US01-16450A-1632	Sequence 1632, Ap
44	39	75.0	376	1	PCT-US01-16450A-1632	Sequence 1632, Ap
45	39	75.0	376	18	US-09-488-725A-2041	Sequence 2041, Ap

ALIGNMENTS

RESULT 1

US-09-763-397A-15  
; Sequence 15, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: the Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-15

Query Match 100.0%; Score 52; DB 21; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GQGHMHG 8

Mon Feb 4 15:23:37 2002

```

Db 1 GQGHMHG 8
|||||
RESULT 2
PCT-US99-26796-59
; Sequence 59, Application PC/TUS9926796
; GENERAL INFORMATION:
; APPLICANT: United States Government as Represented by the
; APPLICANT: Secretary of the Navy
; TITLE OF INVENTION: Chromosome 2 Sequence of Human Malaria Parasite
; TITLE OF INVENTION: Plasmodium falciparum and Proteins of Said
; TITLE OF INVENTION: Chromosome Useful in Anti-malarial Vaccines and
; TITLE OF INVENTION: Diagnostic Reagents
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: PCT/US99/26796
; CURRENT FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 420
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
PCT-US99-26796-59

Query Match 100.0%; Score 52; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 5.4; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 GQGHMHG 8
|||||
Db 206 GQGHMHG 213

RESULT 3
US-09-763-397A-2
; Sequence 2, Application US/09763397A
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis
; APPLICANT: Control and Prevention
; APPLICANT: Lal, Altaf A.
; APPLICANT: Ping Shi, Ya
; APPLICANT: Hasnain, Seyed E.
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci
; FILE REFERENCE: 6395-57049
; CURRENT APPLICATION NUMBER: US/09/763,397A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/097,703
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: PCT / US99/18869
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant DNA/Protein
US-09-763-397A-2

Query Match 100.0%; Score 52; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 6.9; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 GQGHMHG 8
|||||
Db 215 GQGHMHG 222

Query Match 100.0%; Score 43; DB 23; Length 198;
Best Local Similarity 75.0%; Pred. No. 93; Indels 1; Gaps 0;
Matches 6; Conservative 1; Mismatches 1;

QY 1 GQGHMHG 8
|||||
Db 136 GEHGHAG 143

RESULT 5
US-09-902-540-16719
; Sequence 16719, Application US/09902540
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16719
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16719

Query Match 78.8%; Score 41; DB 23; Length 157;
Best Local Similarity 75.0%; Pred. No. 1.5e+02; Indels 1; Gaps 0;
Matches 6; Conservative 1; Mismatches 1;

QY 1 GQGHMHG 8
|||||
Db 142 GEHGHAG 149

RESULT 6
PCT-US01-08631-41146
; Sequence 41146, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
PCT-US01-08631-41146
```

PRIOR APPLICATION NUMBER: 09/649,167  
 PRIOR FILING DATE: 2000-08-23  
 NUMBER OF SEQ ID NOS: 60736  
 SOFTWARE: Custom  
 SEQ ID NO 41146  
 LENGTH: 269  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)...(269)  
 OTHER INFORMATION: Xaa = x or \* as defined in Table 2  
 PCT-US01-08631-41146

Query Match 78.8%; Score 41; DB 1; Length 269;  
 Best Local Similarity 85.7%; Pred. No. 2.5e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QHGHMHG 8  
 | | | | |  
 Db 49 QHGHFHG 55

RESULT 7

US-60-161-932-1547  
 ; Sequence 1547, Application US/60161932  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Venter, J. Craig  
 ; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic  
 ; TITLE OF INVENTION: Acid Sequences, Systems Containin the Nucleic Acid Sequences  
 ; FILE REFERENCE: CL000122  
 ; CURRENT APPLICATION NUMBER: US/60/161,932  
 ; CURRENT FILING DATE: 1999-10-28  
 ; NUMBER OF SEQ ID NOS: 2626  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1547  
 ; LENGTH: 289  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila  
 ; US-60-161-932-1547

Query Match 78.8%; Score 41; DB 24; Length 289;  
 Best Local Similarity 75.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHGHMHG 8  
 | | | | |  
 Db 35 GEHGDHG 42

RESULT 8

US-60-167-217-241  
 ; Sequence 241, Application US/60167217  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Li, Peter W. D.  
 ; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC  
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES  
 ; FILE REFERENCE: CL000152  
 ; CURRENT APPLICATION NUMBER: US/60/167,217  
 ; CURRENT FILING DATE: 1999-11-24  
 ; NUMBER OF SEQ ID NOS: 23195  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 241  
 ; LENGTH: 363  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila  
 ; US-60-167-217-241

Query Match 78.8%; Score 41; DB 24; Length 363;  
 Best Local Similarity 75.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QHGHMHG 8  
 | | | | |  
 Db 123 GHGHSHG 130

RESULT 9

US-60-173-464-228  
 ; Sequence 228, Application US/60173464  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Li, Peter W.D.  
 ; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES  
 ; FILE REFERENCE: CL000173  
 ; CURRENT APPLICATION NUMBER: US/60/173,464  
 ; CURRENT FILING DATE: 1999-12-29  
 ; NUMBER OF SEQ ID NOS: 30269  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 228  
 ; LENGTH: 363  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila  
 ; US-60-173-464-228

Query Match 78.8%; Score 41; DB 24; Length 363;  
 Best Local Similarity 75.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QHGHMHG 8  
 | | | | |  
 Db 123 GHGHSHG 130

RESULT 10

US-60-167-324-1944  
 ; Sequence 1944, Application US/60167324  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Li, Peter W.D.  
 ; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC  
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES  
 ; FILE REFERENCE: CL000155  
 ; CURRENT APPLICATION NUMBER: US/60/167,324  
 ; CURRENT FILING DATE: 1999-11-24  
 ; NUMBER OF SEQ ID NOS: 2272  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1944  
 ; LENGTH: 449  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila  
 ; US-60-167-324-1944

Query Match 78.8%; Score 41; DB 24; Length 449;  
 Best Local Similarity 75.0%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHGHMHG 8  
 | | | | |  
 Db 195 GEHGDHG 202

RESULT 11

US-60-191-637-258  
 ; Sequence 258, Application US/60191637  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Venter, J. Craig  
 ; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING

; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CL000392  
; CURRENT APPLICATION NUMBER: US/60/191,637  
; CURRENT FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 42660  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 258  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-60-191-637-258

Query Match 78.8%; Score 41; DB 24; Length 449;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GQGHMHG 8  
| | | | |  
Db 137 GHGHSHG 144

RESULT 12  
US-60-191-637-26251

; Sequence 26251, Application US/60191637  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING  
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CL000392  
; CURRENT APPLICATION NUMBER: US/60/191,637  
; CURRENT FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 42660  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 26251  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-60-191-637-26251

Query Match 78.8%; Score 41; DB 24; Length 449;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GQGHMHG 8  
| | | | |  
Db 137 GHGHSHG 144

RESULT 13  
US-60-191-637-42246

; Sequence 42246, Application US/60191637  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING  
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CL000392  
; CURRENT APPLICATION NUMBER: US/60/191,637  
; CURRENT FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 42660  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42246  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-60-191-637-42246

Query Match 78.8%; Score 41; DB 24; Length 449;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQGHMHG 8  
| | | | |  
Db 195 GEHGHG 202

RESULT 14

US-60-191-681-213  
; Sequence 213, Application US/60191681  
; GENERAL INFORMATION:  
; APPLICANT: Li, Peter, W.D.  
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CL000390  
; CURRENT APPLICATION NUMBER: US/60/191,681  
; CURRENT FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 30973  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 213  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-60-191-681-213

Query Match 78.8%; Score 41; DB 24; Length 449;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GQGHMHG 8  
| | | | |  
Db 137 GHGHSHG 144

RESULT 15

US-60-191-681-20860  
; Sequence 20860, Application US/60191681  
; GENERAL INFORMATION:  
; APPLICANT: Li, Peter, W.D.  
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CL000390  
; CURRENT APPLICATION NUMBER: US/60/191,681  
; CURRENT FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 30973  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20860  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-60-191-681-20860

Query Match 78.8%; Score 41; DB 24; Length 449;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GQGHMHG 8  
| | | | |  
Db 137 GHGHSHG 144

Search completed: January 29, 2002, 10:56:08  
Job time: 2404 sec

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:14 ; Search time 120.95 Seconds  
(without alignments)  
2.421 Million cell updates/sec

Title: US-09-763-397A-15  
Perfect score: 52  
Sequence: 1 GQGHMHG 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2.6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2.6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2.6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2.6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2.6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2.6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2.6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	42	80.8	137	5	US-09-620-111B-2909 Sequence 2909, Ap
2	42	80.8	406	5	US-09-708-427-50273 Sequence 50273, A
3	41	78.8	294	5	US-09-708-427-1474 Sequence 1474, Ap
4	41	78.8	374	5	US-09-708-427-1473 Sequence 1473, Ap
5	41	78.8	408	5	US-09-708-427-1472 Sequence 1472, Ap
6	39	75.0	366	6	US-10-015-127-13079 Sequence 13079, A
7	39	75.0	376	6	US-10-001-137-29 Sequence 29, Appl
8	39	75.0	378	6	US-10-001-137-59 Sequence 59, Appl
9	38	73.1	168	5	US-09-708-427-71051 Sequence 71051, A
10	37	71.2	330	5	US-09-620-394B-8225 Sequence 8225, Ap
11	37	71.2	309	5	US-09-708-427-36336 Sequence 36336, A
12	37	71.2	344	5	US-09-708-427-36335 Sequence 36335, A
13	37	71.2	389	5	US-09-620-394B-8224 Sequence 8224, Ap
14	37	71.2	398	5	US-09-708-427-36334 Sequence 36334, A
15	37	71.2	453	5	US-09-620-394B-8223 Sequence 8223, Ap
16	36	69.2	121	5	US-09-708-427-60122 Sequence 60122, A
17	36	69.2	125	5	US-09-620-111B-4761 Sequence 4761, Ap
18	36	69.2	313	5	US-09-708-427-53064 Sequence 53064, A
19	36	69.2	318	5	US-09-708-427-53063 Sequence 53063, A
20	36	69.2	400	5	US-09-708-427-53062 Sequence 53062, A
21	35	67.3	76	5	US-09-828-792-706 Sequence 706, App
22	35	67.3	148	5	US-09-708-427-43163 Sequence 43163, A
23	35	67.3	261	5	US-09-708-427-3724 Sequence 3724, Ap
24	35	67.3	296	5	US-09-708-427-3723 Sequence 3723, Ap
25	35	67.3	298	5	US-09-708-427-27822 Sequence 27822, A
26	35	67.3	317	5	US-09-708-427-27821 Sequence 27821, A

ALIGNMENTS

RESULT 1

US-09-620-111B-2909  
; Sequence 2909, Application US/09620111B  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide  
; FILE REFERENCE: 2750-1070P  
; CURRENT APPLICATION NUMBER: US/09/620,111B  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 9298  
; SEQ ID NO 2909  
; LENGTH: 137  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..137  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..137  
; OTHER INFORMATION: Ceres Seq. ID 1326309  
US-09-620-111B-2909

Query Match 80.8%; Score 42; DB 5; Length 137;

Best Local Similarity 75.0%; Pred. No. 1.9;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GQGHMHG 8  
DB 50 GRGHRHG 57

RESULT 2

US-09-708-427-50273  
; Sequence 50273, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 50273  
; LENGTH: 406  
; TYPE: PRT

; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..406  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..406  
; OTHER INFORMATION: Ceres Seq. ID 1924393  
US-09-708-427-50273

Query Match 80.8%; Score 42; DB 5; Length 406;  
Best Local Similarity 75.0%; Pred. No. 6;  
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 1 GQGHMHG 8  
1:|||||  
Db 50 GRGHRHG 57

RESULT 3  
US-09-708-427-1474  
; Sequence 1474, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1474  
; LENGTH: 294  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..294  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..294  
; OTHER INFORMATION: Ceres Seq. ID 1807881  
US-09-708-427-1474

Query Match 78.8%; Score 41; DB 5; Length 294;  
Best Local Similarity 100.0%; Pred. No. 6.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGHMHG 8  
1:|||||  
Db 123 GGHMHG 128

RESULT 4  
US-09-708-427-1473  
; Sequence 1473, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1473  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature

; LOCATION: 1..374  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..374  
; OTHER INFORMATION: Ceres Seq. ID 1807880  
US-09-708-427-1473

Query Match 78.8%; Score 41; DB 5; Length 374;  
Best Local Similarity 100.0%; Pred. No. 8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGHMHG 8  
1:|||||  
Db 203 GGHMHG 208

RESULT 5  
US-09-708-427-1472  
; Sequence 1472, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1472  
; LENGTH: 408  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..408  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..408  
; OTHER INFORMATION: Ceres Seq. ID 1807879  
US-09-708-427-1472

Query Match 78.8%; Score 41; DB 5; Length 408;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGHMHG 8  
1:|||||  
Db 237 GGHMHG 242

RESULT 6  
US-10-015-127-13079  
; Sequence 13079, Application US/10015127  
; GENERAL INFORMATION:  
; APPLICANT: Bower, Stanley G.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Spingomonas elodea genome sequences and uses thereof  
; FILE REFERENCE: 38-10(15806)B  
; CURRENT APPLICATION NUMBER: US/10/015,127  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: US 60/252,455  
; PRIOR FILING DATE: 2000-11-22  
; NUMBER OF SEQ ID NOS: 14357  
; SEQ ID NO 13079  
; LENGTH: 366  
; TYPE: PRT  
; ORGANISM: Spingomonas elodea  
US-10-015-127-13079

Query Match 75.0%; Score 39; DB 6; Length 366;  
Best Local Similarity 62.5%; Pred. No. 17;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 299 GRGHLYG 306

## RESULT 7

US-10-001-137-29  
; Sequence 29, Application US/10001137  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; APPLICANT: Glucksmann, Maria Alexandra  
; TITLE OF INVENTION: 47476, 67210, 49875, 46842, 33201, AND 84234,  
; TITLE OF INVENTION: 83378, 84233, 64708, 85041, AND 84234,  
; TITLE OF INVENTION: NOVEL HUMAN GUANINE NUCLEOTIDE DISSOCIATION STIMULATOR,  
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE, DEAD TYPE HELICASE, CENTAURIN,  
; TITLE OF INVENTION: DEHYDROGENASE/REDUCTASE, AND METAL TRANSPORTER  
; FILE REFERENCE: 10448-119001  
; CURRENT APPLICATION NUMBER: US/10/001,137  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/248,362  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: 60/248,331  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: 60/248,365  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: 60/250,077  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/250,327  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/250,176  
; PRIOR FILING DATE: 2000-11-30  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-001-137-29

Query Match 75.0%; Score 39; DB 6; Length 376;  
Best Local Similarity 75.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 162 GGHGSHG 169

## RESULT 8

US-10-001-137-59  
; Sequence 59, Application US/10001137  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; APPLICANT: Glucksmann, Maria Alexandra  
; TITLE OF INVENTION: 47476, 67210, 49875, 46842, 33201,  
; TITLE OF INVENTION: 83378, 84233, 64708, 85041, AND 84234,  
; TITLE OF INVENTION: NOVEL HUMAN GUANINE NUCLEOTIDE DISSOCIATION STIMULATOR,  
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE, DEAD TYPE HELICASE, CENTAURIN,  
; TITLE OF INVENTION: DEHYDROGENASE/REDUCTASE, AND METAL TRANSPORTER  
; FILE REFERENCE: 10448-119001  
; CURRENT APPLICATION NUMBER: US/10/001,137  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/248,362  
; PRIOR FILING DATE: 2000-11-14

; PRIOR APPLICATION NUMBER: 60/248,331  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: 60/248,365  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: 60/250,077  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/250,327  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/250,176  
; PRIOR FILING DATE: 2000-11-30  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 59  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-001-137-59

Query Match 75.0%; Score 39; DB 6; Length 378;  
Best Local Similarity 75.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 162 GGHGSHG 169

## RESULT 9

US-09-708-427-71051  
; Sequence 71051, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 71051  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..168  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..168  
; OTHER INFORMATION: Ceres Seq. ID 1939634  
US-09-708-427-71051

Query Match 73.1%; Score 38; DB 5; Length 168;  
Best Local Similarity 75.0%; Pred. No. 11;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GQGHMHG 8

Db 31 GAGHGHG 38

## RESULT 10

US-09-620-394B-8225  
; Sequence 8225, Application US/09620394B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 2750-1067P  
; CURRENT APPLICATION NUMBER: US/09/620,394B

; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 9131  
; SEQ ID NO 8225  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..309  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..309  
; OTHER INFORMATION: Ceres Seq. ID 1411807  
US-09-620-394B-8225

Query Match 71.2%; Score 37; DB 5; Length 309;  
Best Local Similarity 83.3%; Pred. No. 30;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QGHMHG 8  
:|||||  
Db 138 EGHVHG 143

RESULT 11  
US-09-708-427-36336  
; Sequence 36336, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36336  
; LENGTH: 330  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..330  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..330  
; OTHER INFORMATION: Ceres Seq. ID 1843032  
US-09-708-427-36336

Query Match 71.2%; Score 37; DB 5; Length 330;  
Best Local Similarity 71.4%; Pred. No. 32;  
Matches 5; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

QY 2 QGHMHG 8  
:|||||  
Db 149 EGHSHG 155

RESULT 12  
US-09-708-427-36335  
; Sequence 36335, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36335

; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..344  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..344  
; OTHER INFORMATION: Ceres Seq. ID 1843031  
US-09-708-427-36335

Query Match 71.2%; Score 37; DB 5; Length 344;  
Best Local Similarity 71.4%; Pred. No. 34;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGHMHG 8  
:|||||  
Db 163 EGHSHG 169

RESULT 13  
US-09-620-394B-8224  
; Sequence 8224, Application US/09620394B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 2750-1067P  
; CURRENT APPLICATION NUMBER: US/09/620,394B  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 9131  
; SEQ ID NO 8224  
; LENGTH: 389  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..389  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..389  
; OTHER INFORMATION: Ceres Seq. ID 1411806  
US-09-620-394B-8224

Query Match 71.2%; Score 37; DB 5; Length 389;  
Best Local Similarity 83.3%; Pred. No. 39;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QGHMHG 8  
:|||||  
Db 218 EGHVHG 223

RESULT 14  
US-09-708-427-36334  
; Sequence 36334, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36334  
; LENGTH: 398  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..398
; OTHER INFORMATION: xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..398
; OTHER INFORMATION: Ceres Seq. ID 1843030
; US-09-708-427-36334
```

```
Query Match 71.2%; Score 37; DB 5; Length 398;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 QHGHMHG 8
Db 217 EHGSHG 223
```

```
RESULT 15
US-09-620-394B-8223
; Sequence 8223, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 8223
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..453
; OTHER INFORMATION: xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..453
; OTHER INFORMATION: Ceres Seq. ID 1411805
; US-09-620-394B-8223
```

```
Query Match 71.2%; Score 37; DB 5; Length 453;
Best Local Similarity 83.3%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3 HGHMHG 8
Db 282 HGHVHG 287
```

```
Search completed: January 29, 2002, 10:58:15
Job time: 2511 sec
```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:06 ; Search time 1760.55 Seconds  
(without alignments)  
1.262 Million cell updates/sec

Title: US-09-763-397A-14  
Perfect score: 41  
Sequence: 1 SNTFINNA 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

#### Database :

Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US060\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	8	21	US-09-763-397A-14
2	41	100.0	272	1	PCT-US99-26796-59
3	41	100.0	350	21	US-09-763-397A-2
4	36	87.8	116	18	US-09-489-039A-13377
5	33	80.5	118	21	US-09-757-028-2046
6	33	80.5	214	12	US-08-827-356-3944
7	33	80.5	214	20	US-09-611-529-7203
8	33	80.5	2010	1	PCT-US99-26796-75
9	32	78.0	76	18	US-09-417-507-29418

10	32	78.0	89	24	US-60-164-763-827
11	32	78.0	89	24	US-60-169-842-4538
12	32	78.0	90	24	US-60-192-587-1563
13	32	78.0	90	24	US-60-194-091-1522
14	32	78.0	127	24	US-60-192-587-995
15	32	78.0	127	24	US-60-194-091-977
16	32	78.0	138	24	US-60-178-308-2852
17	32	78.0	138	24	US-60-192-587-996
18	32	78.0	138	24	US-60-194-091-978
19	32	78.0	141	24	US-60-178-308-2851
20	32	78.0	145	24	US-60-178-308-2850
21	32	78.0	153	24	US-60-170-346-1597
22	32	78.0	158	24	US-60-170-346-1857
23	32	78.0	163	24	US-60-171-480-623
24	32	78.0	166	24	US-60-170-346-1853
25	32	78.0	216	21	US-09-747-155-7
26	32	78.0	216	21	US-09-747-155-55
27	32	78.0	233	24	US-60-324-109-32055
28	32	78.0	259	17	US-09-328-352-7755
29	32	78.0	320	24	US-60-229-515-1187
30	32	78.0	324	1	PCT-US01-00663-30092
31	32	78.0	324	22	US-09-864-761-46627
32	32	78.0	324	24	US-60-236-359-19057
33	32	78.0	341	12	US-08-827-356-4159
34	32	78.0	341	20	US-09-611-529-5799
35	32	78.0	507	24	US-60-324-109-29473
36	32	78.0	648	17	US-09-328-352-5153
37	31	75.6	66	18	US-09-417-507-27537
38	31	75.6	101	17	US-09-391-631-1298
39	31	75.6	101	17	US-09-391-631-4196
40	31	75.6	110	21	US-09-732-210-1205
41	31	75.6	110	24	US-60-169-340-1205
42	31	75.6	133	24	US-60-196-713-3281
43	31	75.6	143	16	US-09-252-691-10602
44	31	75.6	143	16	US-09-252-691C-10602
45	31	75.6	162	24	US-60-192-587-1079

#### ALIGNMENTS

#### RESULT 1

US-09-763-397A-14  
; Sequence 14, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasmain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-14

Query Match 100.0%; Score 41; DB 21; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8

RESULT 4  
US-09-489-039A-13377  
; Sequence 13377, Application US/09489039A  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 13377  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-13377  
  
Query Match 87.8%; Score 36; DB 18; Length 116;  
Best Local Similarity 85.7%; Pred. No. 26;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 NTFINNA 8  
| | | | |  
DB 97 NTFVNA 103  
  
RESULT 5  
US-09-757-028-2046  
; Sequence 2046, Application US/09757028  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PM001  
; CURRENT APPLICATION NUMBER: US/09/757,028  
; CURRENT FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; NUMBER OF SEQ ID NOS: 2660  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2046  
; LENGTH: 118  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (42)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (62)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-757-028-2046  
  
Query Match 80.5%; Score 33; DB 21; Length 118;  
Best Local Similarity 75.0%; Pred. No. 11e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 SNTFINNA 8  
| | | | |  
DB 25 SNTFLNCA 32  
  
RESULT 6  
US-08-827-356-3944  
; Sequence 3944, Application US/08827356  
; GENERAL INFORMATION:  
; APPLICANT: George H. Shimer, Jr.  
; APPLICANT: George H. Miller

DB 1 SNTFINNA 8  
| | | | |  
  
PCT-US99-26796-59  
; Sequence 59, Application PC/TUS9926796  
; GENERAL INFORMATION:  
; APPLICANT: United States Government as Represented by the  
; APPLICANT: Secretary of the Navy  
; TITLE OF INVENTION: Chromosome 2 Sequence of Human Malaria Parasite  
; TITLE OF INVENTION: Plasmodium falciparum and Proteins of Said  
; TITLE OF INVENTION: Chromosome Useful in Anti-malarial Vaccines and  
; TITLE OF INVENTION: Diagnostic Reagents  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: PCT/US99/26796  
; CURRENT FILING DATE: 1999-11-05  
; NUMBER OF SEQ ID NOS: 420  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 59  
; LENGTH: 272  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
PCT-US99-26796-59  
  
Query Match 100.0%; Score 41; DB 1; Length 272;  
Best Local Similarity 100.0%; Pred. No. 6.2; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0;  
  
QY 1 SNTFINNA 8  
| | | | |  
DB 27 SNTFINNA 34  
  
RESULT 3  
US-09-763-397A-2  
; Sequence 2, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Recombinant DNA/Protein  
US-09-763-397A-2  
  
Query Match 100.0%; Score 41; DB 21; Length 350;  
Best Local Similarity 100.0%; Pred. No. 8.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNTFINNA 8  
| | | | |  
DB 207 SNTFINNA 214



APPLICANT: Roberta S. Hare  
APPLICANT: Karen J. Shaw  
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED  
TITLE OF INVENTION: COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 5574  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schering-Plough Corporation  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033-0530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/827,356  
FILING DATE: 01-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/014,477  
FILING DATE: 01-APR-1996  
APPLICATION NUMBER: 60/016,743  
FILING DATE: 02-MAY-1996  
APPLICATION NUMBER: 60/020,016  
FILING DATE: 14-JUN-1996  
INFORMATION FOR SEQ ID NO: 3944:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 214 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Staphylococcus aureus  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1...214  
US-08-827-356-3944

Query Match 80.5%; Score 33; DB 12; Length 214;  
Best Local Similarity 71.4%; Pred. No. 2e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINN 7  
Db 93 ANTEVNN 99

RESULT 7  
US-09-611-529-7203  
Sequence 7203, Application US/09611529  
GENERAL INFORMATION:  
APPLICANT: George H. Shlmer, Jr.  
APPLICANT: George H. Miller  
APPLICANT: Roberta S. Hare  
APPLICANT: Karen J. Shaw  
TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods  
FILE REFERENCE: 1034/IC963US1  
CURRENT APPLICATION NUMBER: US/09/611,529  
CURRENT FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/417,811  
PRIOR FILING DATE: 1999-10-14  
PRIOR APPLICATION NUMBER: US 09/353,718  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: US 09/266,557  
PRIOR FILING DATE: 1999-03-11  
PRIOR APPLICATION NUMBER: US 09/266,556  
PRIOR FILING DATE: 1999-03-11  
PRIOR APPLICATION NUMBER: US 09/266,555  
PRIOR FILING DATE: 1999-03-11

PRIOR APPLICATION NUMBER: US 09/266,542  
PRIOR FILING DATE: 1999-03-11  
PRIOR APPLICATION NUMBER: US 09/266,541  
PRIOR FILING DATE: 1999-03-11  
PRIOR APPLICATION NUMBER: US 09/037,934  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: US 09/036,720  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: US 09/036,338  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: US 09/036,334  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: US 09/036,221  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: US 09/036,137  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: US 09/036,082  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: US 09/036,081  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: US 09/036,079  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: US 09/035,913  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: US 09/035,744  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: US 08/827,356  
PRIOR FILING DATE: 1997-04-01  
PRIOR APPLICATION NUMBER: US 08/831,156  
PRIOR FILING DATE: 1997-04-01  
PRIOR APPLICATION NUMBER: US 60/014,477  
PRIOR FILING DATE: 1996-04-01  
PRIOR APPLICATION NUMBER: US 60/016,743  
PRIOR FILING DATE: 1996-05-02  
PRIOR APPLICATION NUMBER: US 60/020,016  
PRIOR FILING DATE: 1996-06-14  
NUMBER OF SEQ ID NOS: 7451  
SEQ ID NO 7203  
LENGTH: 214  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-611-529-7203

Query Match 80.5%; Score 33; DB 20; Length 214;  
Best Local Similarity 71.4%; Pred. No. 2e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINN 7  
Db 93 ANTEVNN 99

RESULT 8  
PCT-US99-26796-75  
Sequence 75, Application PC/TUS9926796  
GENERAL INFORMATION:  
APPLICANT: United States Government as Represented by the  
APPLICANT: Secretary of the Navy  
TITLE OF INVENTION: Chromosome 2 Sequence of Human Malaria Parasite  
TITLE OF INVENTION: Plasmodium falciparum and Proteins of Said  
TITLE OF INVENTION: Chromosome Useful in Anti-malarial Vaccines and  
TITLE OF INVENTION: Diagnostic Reagents  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: PCT/US99/26796  
CURRENT FILING DATE: 1999-11-05  
NUMBER OF SEQ ID NOS: 420  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 75  
LENGTH: 2010  
TYPE: PRT  
ORGANISM: Plasmodium falciparum  
PCT-US99-26796-75

```
; FILE REFERENCE: CL000162
; CURRENT APPLICATION NUMBER: US/60/169,842
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 5232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4538
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Human
US-60-169-842-4538

Query Match      78.0%; Score 32; DB 24; Length 89;
Best Local Similarity 85.7%; Pred. NO. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINN 7
Db 13 SDTFINN 19

RESULT 12
US-60-192-587-1563 Application US/60192587
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000405
; CURRENT APPLICATION NUMBER: US/60/192,587
; CURRENT FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 1666
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1563
; LENGTH: 90
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(90)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-192-587-1563

Query Match      78.0%; Score 32; DB 24; Length 90;
Best Local Similarity 85.7%; Pred. NO. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINN 7
Db 73 SDTFINN 79

RESULT 13
US-60-194-091-1522
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000416
; CURRENT APPLICATION NUMBER: US/60/194,091
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 1598
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1522
; LENGTH: 90
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
```

```
; FILE REFERENCE: CL000162
; CURRENT APPLICATION NUMBER: US/60/169,842
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 5232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4538
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Human
US-60-169-842-4538

Query Match      78.0%; Score 33; DB 1; Length 2010;
Best Local Similarity 100.0%; Pred. NO. 2.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NTFINN 7
Db 735 NTFINN 740

RESULT 9
US-09-417-507-29418
; Sequence 29418, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; TITLE OF INVENTION: FUMIGATUS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 29418
; LENGTH: 76
; TYPE: PRT
; ORGANISM: A.fumigatus
US-09-417-507-29418

Query Match      78.0%; Score 32; DB 18; Length 76;
Best Local Similarity 85.7%; Pred. NO. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFINN 7
Db 43 SNTFINN 49

RESULT 10
US-60-164-763-827
; Sequence 827, Application US/60164763
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN GPCR PROTEIN, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN GPCR PROTEINS AND USES THEREOF
; FILE REFERENCE: CL000140
; CURRENT APPLICATION NUMBER: US/60/164,763
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 1154
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 827
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Human
US-60-164-763-827

Query Match      78.0%; Score 32; DB 24; Length 89;
Best Local Similarity 85.7%; Pred. NO. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINN 7
Db 13 SDTFINN 19

RESULT 11
US-60-169-842-4538
; Sequence 4538, Application US/60169842
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN GPCR PROTEIN, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN GPCR PROTEINS AND USES THEREOF
```

Search completed: January 29, 2002, 10:56:07  
Job time: 2403 sec

; LOCATION: (1)...(90)  
; OTHER INFORMATION: Xaa - Any Amino Acid  
US-60-194-091-1522

Query Match 78.08; Score 32; DB 24; Length 90;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNTFINN 7  
Db 73 SOTFINN 79  
1:|||||

RESULT 14  
US-60-192-587-995  
; Sequence 995 Application US/60192587  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000405  
; CURRENT APPLICATION NUMBER: US/60/192,587  
; CURRENT FILING DATE: 2000-03-28  
; NUMBER OF SEQ ID NOS: 1666  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 995  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-192-587-995

Query Match 78.08; Score 32; DB 24; Length 127;  
Best Local Similarity 85.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNTFINN 7  
Db 112 SOTFINN 118  
1:|||||

RESULT 15  
US-60-194-091-977  
; Sequence 977 Application US/60194091  
; GENERAL INFORMATION:  
; APPLICANT: Bonazzi, Vivien  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000416  
; CURRENT APPLICATION NUMBER: US/60/194,091  
; CURRENT FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 1598  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 977  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-194-091-977

Query Match 78.08; Score 32; DB 24; Length 127;  
Best Local Similarity 85.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNTFINN 7  
Db 112 SOTFINN 118  
1:|||||



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:14 ; Search time 120.95 seconds  
(without alignments)  
2.421 Million cell updates/sec

Title: US-09-763-397A-14  
Perfect score: 41  
Sequence: 1 SNTFINNA 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	75.6	143	5	US-09-708-427-76264, A
2	31	75.6	152	5	US-09-708-427-76263, A
3	31	75.6	190	5	US-09-708-427-76262, A
4	31	75.6	275	5	US-09-620-394B-4936, Ap
5	31	75.6	329	5	US-09-708-427-54646, A
6	31	75.6	343	5	US-09-708-427-54645, A
7	31	75.6	366	5	US-09-620-394B-4935, Ap
8	31	75.6	374	5	US-09-620-394B-4934, Ap
9	31	75.6	499	5	US-09-881-165-2, Ap
10	31	75.6	621	6	US-10-005-368-41, Ap
11	30	73.2	449	5	US-09-708-427-19898, A
12	30	73.2	502	5	US-09-708-427-19897, A
13	30	73.2	518	5	US-09-708-427-19896, A
14	30	73.2	979	5	US-09-801-368-284, A
15	30	73.2	979	5	US-09-487-558-284, A
16	29	70.7	135	1	PCT-US01-17103-2
17	29	70.7	135	5	US-09-866-582-2, Ap
18	29	70.7	503	5	US-09-801-368-424, A
19	29	70.7	503	5	US-09-487-558-424, A
20	29	70.7	536	5	US-09-816-028A-19, Ap
21	29	70.7	536	5	US-09-816-028A-25, Ap
22	29	70.7	833	7	US-60-337-358-480, Ap
23	28	68.3	66	5	US-09-708-427-79046, A
24	28	68.3	164	5	US-09-708-427-22562, A
25	28	68.3	176	5	US-09-708-427-68823, A
26	28	68.3	189	5	US-09-800-321A-48, Ap

27	28	68.3	208	5	US-09-708-427-68822, A
28	28	68.3	210	5	US-09-708-427-22561, A
29	28	68.3	240	5	US-09-708-427-68821, A
30	28	68.3	244	5	US-09-708-427-17255, A
31	28	68.3	264	5	US-09-708-427-22560, A
32	28	68.3	289	5	US-09-815-242-13111, A
33	28	68.3	292	5	US-09-940-037A-2, Ap
34	28	68.3	292	5	US-09-855-309-2, Ap
35	28	68.3	296	5	US-09-940-037A-23, Ap
36	28	68.3	296	5	US-09-855-309-23, Ap
37	28	68.3	311	5	US-09-779-679-61, Ap
38	28	68.3	311	5	US-09-708-427-17254, A
39	28	68.3	314	5	US-09-886-055-11, Ap
40	28	68.3	314	5	US-09-804-291-11, Ap
41	28	68.3	314	5	US-09-800-321A-8, Ap
42	28	68.3	314	5	US-09-800-321A-10, Ap
43	28	68.3	314	5	US-09-800-321A-12, Ap
44	28	68.3	314	5	US-09-800-321A-14, Ap
45	28	68.3	344	5	US-09-708-427-17253, A

ALIGNMENTS

RESULT 1

US-09-708-427-76264  
; Sequence 76264, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 76264  
; LENGTH: 143  
; TYPE: PPT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..143  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..143  
; OTHER INFORMATION: Ceres Seq. ID 1959859  
US-09-708-427-76264

Query Match 75.6%; Score 31; DB 5; Length 143;  
Best Local Similarity 85.7%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NTFINNA 8  
Db 57 NLFINNA 63

RESULT 2

US-09-708-427-76263  
; Sequence 76263, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 76263  
; LENGTH: 152

```
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..152
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..152
; OTHER INFORMATION: Ceres Seq. ID 1959858
; US-09-708-427-76263

Query Match          75.6%; Score 31; DB 5; Length 152;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NTFINNA 8
Db 66 NLFINNA 72

RESULT 3
US-09-708-427-76262
; Sequence 76262, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76262
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..190
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..190
; OTHER INFORMATION: Ceres Seq. ID 1959857
; US-09-708-427-76262

Query Match          75.6%; Score 31; DB 5; Length 190;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NTFINNA 8
Db 104 NLFINNA 110

RESULT 4
US-09-620-394B-4936
; Sequence 4936, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 4936
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: 1..275
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..275
; OTHER INFORMATION: Ceres Seq. ID 1393687
; US-09-620-394B-4936

Query Match          75.6%; Score 31; DB 5; Length 275;
Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SNTFINN 7
Db 72 SDFINN 78

RESULT 5
US-09-708-427-54646
; Sequence 54646, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54646
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..329
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..329
; OTHER INFORMATION: Ceres Seq. ID 1933688
; US-09-708-427-54646

Query Match          75.6%; Score 31; DB 5; Length 329;
Best Local Similarity 62.5%; Pred. No. 64;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNTFINNA 8
Db 311 SSSPVNNA 318

RESULT 6
US-09-708-427-54645
; Sequence 54645, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54645
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..343
; OTHER INFORMATION: Xaa is any amino acid
```

; NAME/KEY: misc\_feature  
; LOCATION: 1..343  
; OTHER INFORMATION: Ceres Seq. ID 1933687  
US-09-708-427-54645

Query Match 75.6%; Score 31; DB 5; Length 343;  
Best Local Similarity 62.5%; Pred. No. 67;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINNA 8  
||:|:|:|  
Db 325 SSSFVNNA 332

RESULT 7  
US-09-620-394B-4935  
; Sequence 4935, Application US/09620394B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1067P  
; CURRENT APPLICATION NUMBER: US/09/620,394B  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 9131  
; SEQ ID NO 4935  
; LENGTH: 366  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..366  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..366  
; OTHER INFORMATION: Ceres Seq. ID 1393686  
US-09-620-394B-4935

Query Match 75.6%; Score 31; DB 5; Length 366;  
Best Local Similarity 85.7%; Pred. No. 72;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFINN 7  
||:|:|:|  
Db 163 SNDFINN 169

RESULT 8  
US-09-620-394B-4934  
; Sequence 4934, Application US/09620394B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1067P  
; CURRENT APPLICATION NUMBER: US/09/620,394B  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 9131  
; SEQ ID NO 4934  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..374  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..374  
; OTHER INFORMATION: Ceres Seq. ID 1393685

US-09-620-394B-4934

Query Match 75.6%; Score 31; DB 5; Length 374;  
Best Local Similarity 85.7%; Pred. No. 74;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFINN 7  
||:|:|:|  
Db 171 SNDFINN 177

RESULT 9  
US-09-881-165-2  
; Sequence 2, Application US/09881165  
; GENERAL INFORMATION:  
; APPLICANT: HOOD, ELIZABETH  
; APPLICANT: HOWARD, JOHN  
; APPLICANT: BAILEY, MICHELE  
; APPLICANT: GASTEL, FRANS VAN  
; APPLICANT: WANG, HUAMING  
; APPLICANT: WARD, MICHAEL  
; APPLICANT: WOODARD, SUSAN  
; TITLE OF INVENTION: METHOD OF INCREASING RECOVERY OF HETEROLOGOUS ACTIVE  
; FILE REFERENCE: 10032R  
; CURRENT APPLICATION NUMBER: US/09/881,165  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/211,732  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 499  
; TYPE: PRT  
; ORGANISM: Trametes sp.  
US-09-881-165-2

Query Match 75.6%; Score 31; DB 5; Length 499;  
Best Local Similarity 75.0%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNTFINNA 8  
:|:|:|:|  
Db 335 TNEFINNA 342

RESULT 10  
US-10-005-368-41  
; Sequence 41, Application US/10005368  
; GENERAL INFORMATION:  
; APPLICANT: Xun XU et al.  
; TITLE OF INVENTION: PRIMARY NUCLEOTIDE SEQUENCE OF THE  
; TITLE OF INVENTION: SHRIMP WHITE SPOT BACILLIFORM VIRUS (WSBV), DISCOVERY  
; TITLE OF INVENTION: SYSTEMS CONTAINING THIS SEQUENCE AND DETECTION KITS AND  
; TITLE OF INVENTION: ANTIVIRAL TARGETS FOR DETECTION AND CONTROLLING SHRIMP VIRUS  
; TITLE OF INVENTION: OUTBREAK AND SPREAD  
; FILE REFERENCE: CL000895  
; CURRENT APPLICATION NUMBER: US/10/005,368  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 99124717.5  
; PRIOR FILING DATE: 1999-11-24  
; PRIOR APPLICATION NUMBER: PCT/US00/28888  
; PRIOR FILING DATE: 2000-11-08  
; NUMBER OF SEQ ID NOS: 293  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 621  
; TYPE: PRT  
; ORGANISM: SHRIMP  
US-10-005-368-41

Query Match 75.6%; Score 31; DB 6; Length 621;  
Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFINN 7  
Db 263 SSTFVNN 269  
I:|||||

RESULT 11  
US-09-708-427-19898  
; Sequence 19898, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19898  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: misc\_feature  
; LOCATION: 1..449  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..449  
; OTHER INFORMATION: Ceres Seq. ID 1836876  
US-09-708-427-19898

Query Match 73.2%; Score 30; DB 5; Length 449;  
Best Local Similarity 83.3%; Pred. No. 1.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TFINNA 8  
Db 239 TTVNNA 244  
I:|||||

RESULT 12  
US-09-708-427-19897  
; Sequence 19897, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19897  
; LENGTH: 502  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: misc\_feature  
; LOCATION: 1..502  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..502  
; OTHER INFORMATION: Ceres Seq. ID 1836875  
US-09-708-427-19897

Query Match 73.2%; Score 30; DB 5; Length 502;  
Best Local Similarity 83.3%; Pred. No. 1.6e+02;

QY 3 TFINNA 8  
Db 239 TTVNNA 244  
I:|||||

RESULT 13  
US-09-708-427-19896  
; Sequence 19896, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19896  
; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: misc\_feature  
; LOCATION: 1..518  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..518  
; OTHER INFORMATION: Ceres Seq. ID 1836874  
US-09-708-427-19896

Query Match 73.2%; Score 30; DB 5; Length 518;  
Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TFINNA 8  
Db 308 TTVNNA 313  
I:|||||

RESULT 14  
US-09-801-368-284  
; Sequence 284, Application US/09801368  
; GENERAL INFORMATION:  
; APPLICANT: Busby, Robert  
; APPLICANT: Cali, Brian  
; APPLICANT: Hecht, Peter  
; APPLICANT: Holtzman, Doug  
; APPLICANT: Madden, Kevin  
; APPLICANT: Maxon, Mary  
; APPLICANT: Milne, Todd  
; APPLICANT: Norman, Thea  
; APPLICANT: Royer, John  
; APPLICANT: Salama, Sofie  
; APPLICANT: Sherman, Amir  
; APPLICANT: Silva, Jeff  
; APPLICANT: Summers, Eric  
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
; FILE REFERENCE: 109272.147  
; CURRENT APPLICATION NUMBER: US/09/801,368  
; CURRENT FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 09/487,558  
; PRIOR FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US 60/160,587  
; PRIOR FILING DATE: 1999-10-20  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 284  
; LENGTH: 979  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae



US-09-801-368-284

Query Match 73.2%; Score 30; DB 5; Length 979;  
Best Local Similarity 83.3%; Pred. No. 3.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFFIN 6  
Db 668 SNTFVN 673

RESULT 15

US-09-487-558-284  
; Sequence 284, Application US/09487558  
; GENERAL INFORMATION:  
; APPLICANT: Busby, Robert  
; APPLICANT: Call, Brian  
; APPLICANT: Hecht, Peter  
; APPLICANT: Holtzman, Doug  
; APPLICANT: Madden, Kevin  
; APPLICANT: Maxon, Mary  
; APPLICANT: Milne, Todd  
; APPLICANT: Norman, Thea  
; APPLICANT: Royer, John  
; APPLICANT: Salama, Sofie  
; APPLICANT: Sherman, Amir  
; APPLICANT: Silva, Jeff  
; APPLICANT: Summers, Eric  
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production  
; FILE OF INVENTION: In Fungi  
; FILE REFERENCE: 109272.147  
; CURRENT APPLICATION NUMBER: US/09/487,558  
; PRIOR FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US/09/801,368  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 09/487,558  
; PRIOR FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US 60/160,587  
; PRIOR FILING DATE: 1999-10-20  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 284  
; LENGTH: 979  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-487-558-284

Query Match 73.2%; Score 30; DB 5; Length 979;  
Best Local Similarity 83.3%; Pred. No. 3.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNTFFIN 6  
Db 668 SNTFVN 673

Search completed: January 29, 2002, 10:58:14  
Job time: 2510 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:05 ; Search time 1760.55 seconds  
(without alignments)  
2.681 Million cell updates/sec

Title: US-09-763-397A-13  
Perfect score: 88  
Sequence: 1 GISYKVLAKYKDDLE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main: \*  
1: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep: \*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep: \*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep: \*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep: \*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep: \*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep: \*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep: \*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep: \*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep: \*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep: \*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep: \*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep: \*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep: \*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep: \*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep: \*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep: \*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep: \*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep: \*  
24: /cgn2\_6/ptodata/2/paa/US060\_COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	ID	Description
1	88	100.0	17	21	US-09-763-397A-13
2	88	100.0	350	21	US-09-763-397A-2
3	88	100.0	375	21	US-09-710-000-8
4	88	100.0	394	3	US-07-867-768A-2
5	88	100.0	394	3	US-07-867-768A-4
6	88	100.0	394	5	US-08-195-705-2
7	88	100.0	394	5	US-08-195-705-4
8	88	100.0	394	19	US-09-500-376-2
9	88	100.0	394	19	US-09-500-376-3

10	88	100.0	402	19	US-09-500-376-16	Sequence 16, Appl
11	48	54.5	725	18	US-09-417-507-37685	Sequence 37685, A
12	45	51.1	189	21	US-09-758-458-380	Sequence 380, App
13	45	51.1	421	23	US-09-948-933-336	Sequence 336, App
14	45	51.1	498	24	US-60-141-856-469	Sequence 469, App
15	45	51.1	1103	24	US-60-230-445-1214	Sequence 1214, Ap
16	45	51.1	1189	24	US-60-207-583-476	Sequence 476, App
17	45	51.1	2470	24	US-60-167-217-10554	Sequence 10554, A
18	45	51.1	2470	24	US-60-173-464-8530	Sequence 8530, Ap
19	45	51.1	2470	24	US-60-191-637-10565	Sequence 10565, A
20	45	51.1	2470	24	US-60-191-681-8267	Sequence 8267, Ap
21	44	50.0	311	16	US-09-201-228A-397	Sequence 397, App
22	44	50.0	315	16	US-09-248-796-18213	Sequence 18213, A
23	44	50.0	450	20	US-09-686-015-13	Sequence 13, Appl
24	44	50.0	511	1	PCT-US97-00092B-15	Sequence 15, Appl
25	44	50.0	511	1	PCT-US97-22623A-21	Sequence 21, Appl
26	44	50.0	511	15	US-09-134-078-21	Sequence 21, Appl
27	43	48.9	87	12	US-08-827-356-3538	Sequence 3538, Ap
28	43	48.9	87	20	US-09-611-529-4984	Sequence 4984, Ap
29	43	48.9	257	24	US-60-173-464-24041	Sequence 24041, A
30	43	48.9	257	24	US-60-191-637-30547	Sequence 30547, A
31	43	48.9	257	24	US-60-191-681-24373	Sequence 24373, A
32	42	47.7	14	1	PCT-US01-26497-32	Sequence 32, Appl
33	42	47.7	18	1	PCT-US01-26497-14	Sequence 14, Appl
34	42	47.7	62	16	US-09-252-691-10970	Sequence 10970, A
35	42	47.7	62	16	US-09-252-691C-10970	Sequence 10970, A
36	42	47.7	64	24	US-60-324-109-32730	Sequence 32730, A
37	42	47.7	110	19	US-09-543-681A-5117	Sequence 5117, Ap
38	42	47.7	183	21	US-09-758-472-7252	Sequence 7252, Ap
39	42	47.7	219	16	US-09-252-691-11295	Sequence 11295, A
40	42	47.7	219	16	US-09-252-691C-11295	Sequence 11295, A
41	42	47.7	228	21	US-09-758-472-5725	Sequence 5725, Ap
42	42	47.7	296	1	PCT-US00-35017A-778	Sequence 778, App
43	42	47.7	308	1	PCT-US01-08631-53411	Sequence 53411, A
44	42	47.7	377	24	US-60-248-505-801	Sequence 801, App
45	42	47.7	489	1	PCT-US01-08631-45489	Sequence 45489, A

ALIGNMENTS

RESULT 1  
US-09-763-397A-13  
; Sequence 13, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-13

Query Match 100.0%; Score 88; DB 21; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.9e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GISYKVLAKYKDDLE 17

Mon Feb 4 15:23:36 2002

us-09-763-397a-13.rapm

Db 1 GISYKVLAKYKDDLE 17  
|||||

RESULT 2  
US-09-763-397A-2

; Sequence 2, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Recombinant DNA/Protein  
US-09-763-397A-2

Query Match 100.0%; Score 88; DB 21; Length 350;

Best Local Similarity 100.0%; Pred. No. 1.6e-05; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17  
|||||

Db 334 GISYKVLAKYKDDLE 350

RESULT 3

US-09-710-000-8  
; Sequence 8, Application US/09710000  
; GENERAL INFORMATION:  
; APPLICANT: Hui, George, S.N.  
; APPLICANT: Ho, Walter K.K.  
; APPLICANT: Lap-Yin, Pang  
; TITLE OF INVENTION: Malaria Vaccine  
; FILE REFERENCE: 23461-2001100  
; CURRENT APPLICATION NUMBER: US/09/710,000  
; CURRENT FILING DATE: 2000-11-10  
; PRIOR APPLICATION NUMBER: 60/226,861  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/165,178  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/168,327  
; PRIOR FILING DATE: 1999-12-11  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of Pfmsp-142  
US-09-710-000-8

Query Match 100.0%; Score 88; DB 21; Length 375;

Best Local Similarity 100.0%; Pred. No. 1.7e-05; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17  
|||||

Db 119 GISYKVLAKYKDDLE 135

RESULT 4

US-07-867-768A-2  
; Sequence 2, Application US/07867768A  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Sandra  
; APPLICANT: Hui, George  
; APPLICANT: Barr, Philip  
; APPLICANT: Gibson, Helen  
; TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMODIUM  
; TITLE OF INVENTION: FALCIPARUM VACCINE  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Davis Hoxie Faithfull Hapgood  
; STREET: 45 Rockefeller Pl.  
; CITY: N.Y.  
; STATE: N.Y.  
; COUNTRY: USA  
; ZIP: 10111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/867,768A  
; FILING DATE: 19920413  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jacobs, Seth H  
; REGISTRATION NUMBER: 32140  
; REFERENCE/DOCKET NUMBER: 11880A3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-757-2200  
; TELEFAX: 212-586-1461  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 394 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: C-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Plasmodium falciparum  
; STRAIN: falciparum uganda palo alto (FUP)  
US-07-867-768A-2

Query Match 100.0%; Score 88; DB 3; Length 394;

Best Local Similarity 100.0%; Pred. No. 1.9e-05; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKVLAKYKDDLE 17  
|||||

Db 118 GISYKVLAKYKDDLE 134

RESULT 5

US-07-867-768A-4  
; Sequence 4, Application US/07867768A  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Sandra  
; APPLICANT: Hui, George  
; APPLICANT: Barr, Philip  
; APPLICANT: Gibson, Helen

```

; TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMIDIUM
; TITLE OF INVENTION: FALCIPARUM VACCINE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Davis Hoxie Faithfull Hapgood
; STREET: 45 Rockefeller Pl.
; CITY: New York
; STATE: N.Y.
; COUNTRY: USA
; ZIP: 10111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,768A
; FILING DATE: 19920413
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Jacobs, Seth H
; REGISTRATION NUMBER: 32140
; REFERENCE/DOCKET NUMBER: 11880A3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-757-2200
; TELEFAX: 212-586-1461
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; STRAIN: MAD
; US-07-867-768A-4

```

Query Match 100.0%; Score 88; DB 3; Length 394;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GISYKVLAKYKDDLE 17
Db 118 GISYKVLAKYKDDLE 134

```

```

RESULT 6
US-08-195-705-2
; Sequence 2, Application US/08195705
; GENERAL INFORMATION:
; APPLICANT: Chang, Sandra
; APPLICANT: Hui, George
; APPLICANT: Barr, Philip
; APPLICANT: Gibson, Helen
; TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMIDIUM
; TITLE OF INVENTION: FALCIPARUM VACCINE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Davis Hoxie Faithfull Hapgood
; STREET: 45 Rockefeller Pl.
; CITY: New York
; STATE: N.Y.
; COUNTRY: USA
; ZIP: 10111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195,705
; FILING DATE: 14-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jacobs, Seth H
; REGISTRATION NUMBER: 32140
; REFERENCE/DOCKET NUMBER: 11880A3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-757-2200
; TELEFAX: 212-586-1461
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; STRAIN: falciparum uganda palo alto (FUP)
; US-08-195-705-2

```

Query Match 100.0%; Score 88; DB 5; Length 394;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GISYKVLAKYKDDLE 17
Db 118 GISYKVLAKYKDDLE 134

```

```

RESULT 7
US-08-195-705-4
; Sequence 4, Application US/08195705
; GENERAL INFORMATION:
; APPLICANT: Chang, Sandra
; APPLICANT: Hui, George
; APPLICANT: Barr, Philip
; APPLICANT: Gibson, Helen
; TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMIDIUM
; TITLE OF INVENTION: FALCIPARUM VACCINE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Davis Hoxie Faithfull Hapgood
; STREET: 45 Rockefeller Pl.
; CITY: New York
; STATE: N.Y.
; COUNTRY: USA
; ZIP: 10111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195,705
; FILING DATE: 14-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jacobs, Seth H
; REGISTRATION NUMBER: 32140
; REFERENCE/DOCKET NUMBER: 11880A3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-757-2200
; TELEFAX: 212-586-1461
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids

```

```

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; STRAIN: MAD
; US-08-195-705-4

Query Match          100.0%; Score 88; DB 5; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYVEKVLAKYKDDLE 17
   |||||
Db 118 GISYVEKVLAKYKDDLE 134

RESULT 8
US-09-500-376-2
; Sequence 2, Application US/09500376
; GENERAL INFORMATION:
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Baculovirus Produced Plasmodium Falciparum Vaccine
; FILE REFERENCE: A-67984
; CURRENT APPLICATION NUMBER: US/09/500,376
; CURRENT FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 08/195,705
; PRIOR FILING DATE: 1994-02-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 394
; TYPE: PRT
; ORGANISM: plasmodium falciparum
US-09-500-376-2

Query Match          100.0%; Score 88; DB 19; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYVEKVLAKYKDDLE 17
   |||||
Db 118 GISYVEKVLAKYKDDLE 134

RESULT 9
US-09-500-376-3
; Sequence 3, Application US/09500376
; GENERAL INFORMATION:
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Baculovirus Produced Plasmodium Falciparum Vaccine
; FILE REFERENCE: A-67984
; CURRENT APPLICATION NUMBER: US/09/500,376
; CURRENT FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 08/195,705
; PRIOR FILING DATE: 1994-02-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 394
; TYPE: PRT
; ORGANISM: plasmodium falciparum
US-09-500-376-3

Query Match          100.0%; Score 88; DB 19; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYVEKVLAKYKDDLE 17
   |||||
Db 118 GISYVEKVLAKYKDDLE 134

RESULT 10
US-09-500-376-16
; Sequence 16, Application US/09500376
; GENERAL INFORMATION:
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Baculovirus Produced Plasmodium Falciparum Vaccine
; FILE REFERENCE: A-67984
; CURRENT APPLICATION NUMBER: US/09/500,376
; CURRENT FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 08/195,705
; PRIOR FILING DATE: 1994-02-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 402
; TYPE: PRT
; ORGANISM: plasmodium falciparum
US-09-500-376-16

Query Match          100.0%; Score 88; DB 19; Length 402;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYVEKVLAKYKDDLE 17
   |||||
Db 139 GISYVEKVLAKYKDDLE 155

RESULT 11
US-09-417-507-37685
; Sequence 37685, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PATH99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 37685
; LENGTH: 725
; TYPE: PRT
; ORGANISM: A.fumigatus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (531)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are un
US-09-417-507-37685

Query Match          54.5%; Score 48; DB 18; Length 725;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 YEKVLAKYKDDL 16
   |||
Db 170 YEEFVAKYKDDL 181

RESULT 12
US-09-758-458-380
; Sequence 380, Application US/09758458
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM041
```

```

: CURRENT APPLICATION NUMBER: US/09/758,458
:
: PRIOR FILING DATE: 2001-01-11
: PRIOR APPLICATION NUMBER: 60/179,065
: PRIOR FILING DATE: 2000-01-31
: PRIOR APPLICATION NUMBER: 60/180,628
: PRIOR FILING DATE: 2000-02-04
: NUMBER OF SEQ ID NOS: 520
: SOFTWARE: PatentIn ver. 2.0
:
: SEQ ID NO 380
: LENGTH: 189
: TYPE: PRT
:
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (28) .
: OTHER INFORMATION: Xaa equals any of the
: NAME/KEY: SITE
: LOCATION: (60)
: OTHER INFORMATION: Xaa equals any of the
: NAME/KEY: SITE
: LOCATION: (96)
: OTHER INFORMATION: Xaa equals any of the
: NAME/KEY: SITE
: LOCATION: (113)
: OTHER INFORMATION: Xaa equals any of the
: NAME/KEY: SITE
: LOCATION: (114)
: OTHER INFORMATION: Xaa equals any of the
: US-09-758-458-380

```

Query Match 51.1%; Score 45; DB 21; Length 189;  
Best Local Similarity 41.2%; Pred. NO. 69;  
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 GISYKVLAKYKDDE 17  
        ||||| : | : : : | :  
Db     68 GIDYDRNLALFEELD 84
```

```

RESULT 13
US-09-948-933-336
; Sequence 336, Application US/09948933
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH CARDIOVASCULAR DISEASE AND LIPID DISORDERS, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL000787
; CURRENT APPLICATION NUMBER: US/09/948,933
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,399
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 6404
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 336
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Human
US-09-948-933-336

```

Query Match 51.1%; Score 45; DB 23; Length 421;  
Best Local Similarity 41.2%; Pred. No. 1.8e+02;  
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

```
Qy 1 GISYEVVLAKYKDDLE 17
    || ||: || : : : | :
Db 92 GIDYDRNLALFEELD 108
```

RESULT 14  
US-60-141-856-469

```

: Sequence 459, Application US/60141856
:
: GENERAL INFORMATION:
: APPLICANT: Kurlavage, Anthony
: TITLE OF INVENTION: ISOLATED KINASE PROTEINS, NUCLEIC ACID
: TITLE OF INVENTION: MOLECULES ENCODING KINASE PROTEINS, AND USES THEREOF
: FILE REFERENCE: CL000044
: CURRENT APPLICATION NUMBER: US/60/141,856
: CURRENT FILING DATE: 1999-07-01
: NUMBER OF SEQ ID NOS: 617
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 459
: LENGTH: 498
: TYPE: PRT
: ORGANISM: Drosophila
:
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)...(498)
: OTHER INFORMATION: xaa = Any Amino Acid
US-60-141-856-469

```

Query Match	51.1%	Score 45	DB 24	Length 498
Best Local Similarity	60.0%	Pred. No.	2.3e+02	
Matches	9	Conservative	1	Mismatches 5
				Indels 0
				Gaps 0

Qy	3	SYE	KVL	AKY	KDD	LE	17
				:			
Db	241	SYX	RAV	LA	VH	HD	255
				:			

```

RESULT 15
US-60-230-445-1214
; Sequence 1214, Application US/60230445
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000765
; CURRENT APPLICATION NUMBER: US/60/230,445
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 3051
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1214
; LENGTH: 1103
; TYPE: PRT
; ORGANISM: HUMAN
; US-60-230-445-1214

```

Query Match	51.1%	Score 45	DB 24	Length 1103
Best Local Similarity	41.2%	Pred. No. 6e+02		
Matches	7	Conservative	7	Mismatches 3
				Indels 0
				Gaps 0

```

QY      1 GISYKVLAKYKDDLE 17
      ||||| : : : :
Db     58 GIDYDRNLALFEELD 74

```

Search completed: January 29, 2002, 10:56:06  
Job time: 2402 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:13 ; Search time 120.95 seconds  
(without alignments)  
5.144 Million cell updates/sec

Title: US-09-763-397A-13  
Perfect score: 88  
Sequence: 1 GISYKEKVLAKYKDDLE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2.6/ptodata/2/paa/PCT\_NEW\_COMB.pcp.\*  
2: /cgn2.6/ptodata/2/paa/US06\_NEW\_COMB.pcp.\*  
3: /cgn2.6/ptodata/2/paa/US07\_NEW\_COMB.pcp.\*  
4: /cgn2.6/ptodata/2/paa/US08\_NEW\_COMB.pcp.\*  
5: /cgn2.6/ptodata/2/paa/US09\_NEW\_COMB.pcp.\*  
6: /cgn2.6/ptodata/2/paa/US10\_NEW\_COMB.pcp.\*  
7: /cgn2.6/ptodata/2/paa/US60\_NEW\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	394	5	US-09-978-756-3
2	42	47.7	14	5	US-09-645-454-32
3	42	47.7	18	5	US-09-645-454-14
4	42	47.7	280	5	US-09-708-427-51280
5	42	47.7	293	5	US-09-708-427-51279
6	42	47.7	306	5	US-09-708-427-51278
7	41	46.6	376	5	US-09-978-756-2
8	41	46.6	594	5	US-09-117-415B-2
9	41	46.6	613	5	US-09-117-415B-22
10	41	46.6	631	5	US-09-117-415B-18
11	41	46.6	631	5	US-09-117-415B-20
12	41	46.6	649	5	US-09-117-415B-16
13	41	46.6	1602	5	US-09-269-874A-7
14	41	46.6	1621	5	US-09-269-874A-5
15	41	46.6	1639	5	US-09-269-874A-3
16	40	45.5	164	5	US-09-620-394B-2020
17	40	45.5	201	5	US-09-620-394B-2019
18	40	45.5	243	5	US-09-968-521-70
19	40	45.5	264	5	US-09-673-481A-4
20	40	45.5	272	5	US-09-966-521-2
21	40	45.5	286	5	US-09-673-481A-20
22	40	45.5	298	5	US-09-620-394B-2018
23	40	45.5	309	5	US-09-815-242-5459
24	40	45.5	337	5	US-09-815-242-12254
25	40	45.5	562	5	US-09-897-516-6447
26	40	45.5	785	5	US-09-897-516-7123

27	40	45.5	996	5	US-09-815-242-5251	Sequence 5251, Ap
28	40	45.5	1009	5	US-09-815-242-12141	Sequence 12141, A
29	39	44.3	65	5	US-09-897-516-6564	Sequence 6564, Ap
30	39	44.3	426	5	US-09-815-242-11298	Sequence 11298, A
31	39	44.3	426	5	US-09-815-242-11476	Sequence 11476, A
32	39	44.3	581	5	US-09-708-427-10298	Sequence 10298, A
33	39	44.3	640	6	US-10-015-127-11286	Sequence 11286, A
34	39	44.3	875	5	US-09-708-427-10440	Sequence 10440, A
35	39	44.3	907	5	US-09-708-427-10439	Sequence 10439, A
36	39	44.3	925	5	US-09-708-427-10438	Sequence 10438, A
37	38	43.2	287	5	US-09-708-427-32720	Sequence 32720, A
38	38	43.2	291	6	US-10-015-127-12474	Sequence 12474, A
39	38	43.2	309	5	US-09-613-092A-2	Sequence 32719, A
40	38	43.2	342	5	US-09-708-427-32719	Sequence 32719, A
41	38	43.2	381	5	US-09-815-242-11280	Sequence 11280, A
42	38	43.2	402	5	US-09-815-242-11534	Sequence 11534, A
43	38	43.2	406	5	US-09-815-242-13392	Sequence 13392, A
44	38	43.2	435	5	US-09-708-427-32718	Sequence 32718, A
45	38	43.2	474	5	US-09-815-242-10270	Sequence 10270, A

ALIGNMENTS

RESULT 1  
US-09-978-756-3  
; Sequence 3, Application US/09978756  
; GENERAL INFORMATION:  
; APPLICANT: Holder, Anthony  
; APPLICANT: Birdsell, Berry  
; APPLICANT: Feeney, James  
; APPLICANT: Morgan, William  
; APPLICANT: Syed, Shabih  
; TITLE OF INVENTION: Malaria Vaccine  
; FILE REFERENCE: 18396/1005  
; CURRENT APPLICATION NUMBER: US/09/978,756  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: PCT/GB00/01558  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 09/311,817  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: 2,271,451  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 9909072.2  
; PRIOR FILING DATE: 1999-04-20  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-978-756-3

Query Match 100.0%; Score 88; DB 5; Length 394;  
Best Local Similarity 100.0%; Pred. No. 4e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GISYKEKVLAKYKDDLE 17  
DB 118 GISYKEKVLAKYKDDLE 134

RESULT 2  
US-09-645-454-32  
; Sequence 32, Application US/09645454  
; GENERAL INFORMATION:  
; APPLICANT: Fogelman, Alan  
; APPLICANT: Anantharamaiah, Gattadahlalli  
; APPLICANT: Navab, Mohamad  
; TITLE OF INVENTION: Orally Administered Peptides to Ameliorate Atherosclerosis  
; FILE REFERENCE: 407T-911200US  
; CURRENT APPLICATION NUMBER: US/09/645,454

```

; CURRENT FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: L is blocked with an acetyl
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: F is blocked with an amide
; US-09-645-454-32

```

Query Match 47.7%; Score 42; DB 5; Length 14;  
 Best Local Similarity 63.6%; Pred. No. 1.1;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 YYEKVLAKYKD 14  
 :||||| :|:  
 Db 2 FYEKVLEKFE 12

```

RESULT 3
US-09-645-454-14
; Sequence 14, Application US/09645454
; GENERAL INFORMATION:
; APPLICANT: Fogelman, Alan
; APPLICANT: Anantharamiah, Gattadahalli
; TITLE OF INVENTION: Orally Administered Peptides to Ameliorate Atherosclerosis
; FILE REFERENCE: 407T-911200US
; CURRENT APPLICATION NUMBER: US/09/645,454
; CURRENT FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: E is blocked with an acetyl
; NAME/KEY: misc_feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: F is blocked with an amide
; US-09-645-454-14

```

Query Match 47.7%; Score 42; DB 5; Length 18;  
 Best Local Similarity 63.6%; Pred. No. 1.5;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 YYEKVLAKYKD 14  
 :||||| :|:  
 Db 6 FYEKVLEKFE 16

```

RESULT 4
US-09-708-427-51280
; Sequence 51280, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1

```

```

; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51280
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..280
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..280
; OTHER INFORMATION: Ceres Seq. ID 1926459
; US-09-708-427-51280

```

Query Match 47.7%; Score 42; DB 5; Length 280;  
 Best Local Similarity 58.3%; Pred. No. 33;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 EKVLAKYKDDLE 17  
 |::||:||||:  
 Db 265 EQIKAKFKDDLD 276

```

RESULT 5
US-09-708-427-51279
; Sequence 51279, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51279
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..293
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..293
; OTHER INFORMATION: Ceres Seq. ID 1926458
; US-09-708-427-51279

```

Query Match 47.7%; Score 42; DB 5; Length 293;  
 Best Local Similarity 58.3%; Pred. No. 35;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 EKVLAKYKDDLE 17  
 |::||:||||:  
 Db 278 EQIKAKFKDDLD 289

```

RESULT 6
US-09-708-427-51278
; Sequence 51278, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1

```

; SEQ ID NO 51278  
; LENGTH: 306  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..306  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..306  
; OTHER INFORMATION: Ceres Seq. ID 1926457  
US-09-708-427-51278

Query Match 47.7%; Score 42; DB 5; Length 306;  
Best Local Similarity 58.3%; Pred. No. 37;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 EKVLAKYKDDLE 17  
I::||:||||:  
Db 291 EQIAKAKFKDDLD 302

RESULT 7  
US-09-978-756-2  
; Sequence 2, Application US/09978756  
; GENERAL INFORMATION:  
; APPLICANT: Holder, Anthony  
; APPLICANT: Birdsall, Berry  
; APPLICANT: Feeney, James  
; APPLICANT: Morgan, William  
; APPLICANT: Syed, Shabih  
; TITLE OF INVENTION: Malaria Vaccine  
; FILE REFERENCE: 18396/1005  
; CURRENT APPLICATION NUMBER: US/09/978,756  
; PRIOR FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: PCT/GB00/01558  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 09/311,817  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: 2,271,451  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 9909072.2  
; PRIOR FILING DATE: 1999-04-20  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-978-756-2

Query Match 46.6%; Score 41; DB 5; Length 376;  
Best Local Similarity 43.8%; Pred. No. 66;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 ISYKVLAKYKDDLE 17  
I::||:||||:  
Db 120 LGYKILSEKYSKDDLD 135

RESULT 8  
US-09-117-415B-2  
; Sequence 2, Application US/09117415B  
; GENERAL INFORMATION:  
; APPLICANT: Davidson, Eugene  
; APPLICANT: Yang, Shutong  
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
; OF A MSAL PEPTIDE  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC

; STREET: 714 Colorado Avenue  
; CITY: Bridgeport  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 10017  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPad (ASCII)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/117,415B  
; FILING DATE: 29-Jul-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coleman, Henry D.  
; REGISTRATION NUMBER: 32,559  
; REFERENCE/DOCKET NUMBER: R12-030  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 679-0090  
; TELEFAX: (212) 679-9121  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 594 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-117-415B-2

Query Match 46.6%; Score 41; DB 5; Length 594;  
Best Local Similarity 43.8%; Pred. No. 1,1e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 ISYKVLAKYKDDLE 17  
I::||:||||:  
Db 338 LGYKILSEKYSKDDLD 353

RESULT 9  
US-09-117-415B-22  
; Sequence 22, Application US/09117415B  
; GENERAL INFORMATION:  
; APPLICANT: Davidson, Eugene  
; APPLICANT: Yang, Shutong  
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
; OF A MSAL PEPTIDE  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
; STREET: 714 Colorado Avenue  
; CITY: Bridgeport  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 10017  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPad (ASCII)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/117,415B  
; FILING DATE: 29-Jul-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coleman, Henry D.  
; REGISTRATION NUMBER: 32,559  
; REFERENCE/DOCKET NUMBER: R12-030  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 679-0090  
; TELEFAX: (212) 679-9121  
; INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:  
LENGTH: 613 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-117-415B-22

Query Match 46.6%; Score 41; DB 5; Length 613;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 ISYKVLAKYKDDLE 17  
: ||| : ||| |||  
Db 375 LGYKILSEKYSDDL 390

RESULT 10  
US-09-117-415B-18  
; Sequence 18, Application US/09117415B  
; GENERAL INFORMATION:  
; APPLICANT: Davidson, Eugene  
; Yang, Shutong  
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
; of a MSAL Peptide

NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
STREET: 714 Colorado Avenue  
CITY: Bridgeport  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 10017

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPad (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/117,415B  
FILING DATE: 29-Jul-1998  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Coleman, Henry D.  
REGISTRATION NUMBER: 32,559  
REFERENCE/DOCKET NUMBER: R12-030  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 679-0090  
TELEFAX: (212) 679-9121

INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 631 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-117-415B-18

Query Match 46.6%; Score 41; DB 5; Length 631;  
Best Local Similarity 43.8%; Pred. No. 1.2e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 ISYKVLAKYKDDLE 17  
: ||| : ||| |||  
Db 393 LGYKILSEKYSDDL 408

RESULT 11  
US-09-117-415B-20  
; Sequence 20, Application US/09117415B  
; GENERAL INFORMATION:

APPLICANT: Davidson, Eugene  
Yang, Shutong  
TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
of a MSAL Peptide  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
STREET: 714 Colorado Avenue  
CITY: Bridgeport  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 10017

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPad (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/117,415B  
FILING DATE: 29-Jul-1998  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Coleman, Henry D.  
REGISTRATION NUMBER: 32,559  
REFERENCE/DOCKET NUMBER: R12-030  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 679-0090  
TELEFAX: (212) 679-9121

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 631 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-117-415B-20

Query Match 46.6%; Score 41; DB 5; Length 631;  
Best Local Similarity 43.8%; Pred. No. 1.2e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 ISYKVLAKYKDDLE 17  
: ||| : ||| |||  
Db 375 LGYKILSEKYSDDL 390

RESULT 12  
US-09-117-415B-16  
; Sequence 16, Application US/09117415B  
; GENERAL INFORMATION:  
; APPLICANT: Davidson, Eugene  
; Yang, Shutong  
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
; of a MSAL Peptide

NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
STREET: 714 Colorado Avenue  
CITY: Bridgeport  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 10017

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPad (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/117,415B  
FILING DATE: 29-Jul-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

; NAME: Coleman, Henry D.  
; REGISTRATION NUMBER: 32,559  
; REFERENCE/DOCKET NUMBER: R12-030  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 679-0090  
; TELEFAX: (212) 679-9121  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 649 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-117-415B-16

Query Match 46.6%; Score 41; DB 5; Length 649;  
Best Local Similarity 43.8%; Pred. No. 1.2e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 2 ISYYEKVLAKYKDDLE 17  
: ||: : ||| ||:  
Db 393 LGYKILSEKYSDDL 408

RESULT 13  
US-09-269-874A-7  
; SEQUENCE 7, Application US/09269874A  
; GENERAL INFORMATION:  
; APPLICANT: Bujard, Hermann  
; TITLE OF INVENTION: Recombinant Process for Preparing a  
; FILE REFERENCE: GRUE-003  
; CURRENT APPLICATION NUMBER: US/09/269,874A  
; CURRENT FILING DATE: 1999-08-02  
; PRIOR APPLICATION NUMBER: PCT/EP97/05441  
; PRIOR FILING DATE: 1997-10-02  
; PRIOR APPLICATION NUMBER: DE 19640817.2  
; PRIOR FILING DATE: 1996-10-02  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 1602  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-269-874A-7

Query Match 46.6%; Score 41; DB 5; Length 1602;  
Best Local Similarity 43.8%; Pred. No. 3.4e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 2 ISYYEKVLAKYKDDLE 17  
: ||: : ||| ||:  
Db 1364 LGYKILSEKYSDDL 1379

RESULT 14  
US-09-269-874A-5  
; SEQUENCE 5, Application US/09269874A  
; GENERAL INFORMATION:  
; APPLICANT: Bujard, Hermann  
; TITLE OF INVENTION: Recombinant Process for Preparing a  
; FILE REFERENCE: GRUE-003  
; CURRENT APPLICATION NUMBER: US/09/269,874A  
; CURRENT FILING DATE: 1999-08-02  
; PRIOR APPLICATION NUMBER: PCT/EP97/05441  
; PRIOR FILING DATE: 1997-10-02  
; PRIOR APPLICATION NUMBER: DE 19640817.2  
; PRIOR FILING DATE: 1996-10-02  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5  
; LENGTH: 1621  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-269-874A-5

Query Match 46.6%; Score 41; DB 5; Length 1621;  
Best Local Similarity 43.8%; Pred. No. 3.5e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 2 ISYYEKVLAKYKDDLE 17  
: ||: : ||| ||:  
Db 1383 LGYKILSEKYSDDL 1398

RESULT 15  
US-09-269-874A-3  
; SEQUENCE 3, Application US/09269874A  
; GENERAL INFORMATION:  
; APPLICANT: Bujard, Hermann  
; TITLE OF INVENTION: Recombinant Process for Preparing a  
; FILE REFERENCE: GRUE-003  
; CURRENT APPLICATION NUMBER: US/09/269,874A  
; CURRENT FILING DATE: 1999-08-02  
; PRIOR APPLICATION NUMBER: PCT/EP97/05441  
; PRIOR FILING DATE: 1997-10-02  
; PRIOR APPLICATION NUMBER: DE 19640817.2  
; PRIOR FILING DATE: 1996-10-02  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1639  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-269-874A-3

Query Match 46.6%; Score 41; DB 5; Length 1639;  
Best Local Similarity 43.8%; Pred. No. 3.5e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 2 ISYYEKVLAKYKDDLE 17  
: ||: : ||| ||:  
Db 1383 LGYKILSEKYSDDL 1398

Search completed: January 29, 2002, 10:58:14  
Job time: 2510 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:05 ; Search time 1760.55 Seconds  
(without alignments)  
2.996 Million cell updates/sec

Title: US-09-763-397A-12

Perfect score: 106

Sequence: 1 EDGSGNGKKTCTKTPDS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main.\*

- 1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US060\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	106	100.0	19	21	US-09-763-397A-12
2	106	100.0	95	15	US-09-125-031-2
3	106	100.0	95	15	US-09-125-031A-2
4	106	100.0	95	15	US-09-125-031B-2
5	106	100.0	95	15	US-09-134-333-2
6	106	100.0	96	17	US-09-311-817-1
7	106	100.0	108	15	US-09-125-031-10
8	106	100.0	108	15	US-09-125-031A-10
9	106	100.0	108	15	US-09-125-031B-10
10	106	100.0	108	15	US-09-125-031B-10
11	106	100.0	108	15	US-09-125-031B-10
12	106	100.0	108	15	US-09-125-031B-10
13	106	100.0	108	15	US-09-125-031B-10
14	106	100.0	108	15	US-09-125-031B-10
15	106	100.0	108	15	US-09-125-031B-10
16	106	100.0	108	15	US-09-125-031B-10
17	106	100.0	108	15	US-09-125-031B-10
18	106	100.0	108	15	US-09-125-031B-10
19	106	100.0	108	15	US-09-125-031B-10
20	106	100.0	108	15	US-09-125-031B-10
21	106	100.0	108	15	US-09-125-031B-10
22	106	100.0	108	15	US-09-125-031B-10
23	106	100.0	108	15	US-09-125-031B-10
24	106	100.0	108	15	US-09-125-031B-10
25	106	100.0	108	15	US-09-125-031B-10
26	106	100.0	108	15	US-09-125-031B-10
27	106	100.0	108	15	US-09-125-031B-10
28	106	100.0	108	15	US-09-125-031B-10
29	106	100.0	108	15	US-09-125-031B-10
30	106	100.0	108	15	US-09-125-031B-10
31	106	100.0	108	15	US-09-125-031B-10
32	106	100.0	108	15	US-09-125-031B-10
33	106	100.0	108	15	US-09-125-031B-10
34	106	100.0	108	15	US-09-125-031B-10
35	106	100.0	108	15	US-09-125-031B-10
36	106	100.0	108	15	US-09-125-031B-10
37	106	100.0	108	15	US-09-125-031B-10
38	106	100.0	108	15	US-09-125-031B-10
39	106	100.0	108	15	US-09-125-031B-10
40	106	100.0	108	15	US-09-125-031B-10
41	106	100.0	108	15	US-09-125-031B-10
42	106	100.0	108	15	US-09-125-031B-10
43	106	100.0	108	15	US-09-125-031B-10
44	106	100.0	108	15	US-09-125-031B-10
45	106	100.0	108	15	US-09-125-031B-10

10	106	100.0	108	15	US-09-134-333-10	Sequence 10, Appl
11	106	100.0	116	15	US-09-125-031-5	Sequence 5, Appl
12	106	100.0	116	15	US-09-125-031A-5	Sequence 5, Appl
13	106	100.0	116	15	US-09-125-031B-5	Sequence 5, Appl
14	106	100.0	116	15	US-09-134-333-5	Sequence 5, Appl
15	106	100.0	127	15	US-09-125-031-8	Sequence 8, Appl
16	106	100.0	127	15	US-09-125-031A-8	Sequence 8, Appl
17	106	100.0	127	15	US-09-125-031B-8	Sequence 8, Appl
18	106	100.0	127	15	US-09-134-333-8	Sequence 8, Appl
19	106	100.0	350	21	US-09-763-397A-2	Sequence 2, Appl
20	106	100.0	355	15	US-09-175-683-11	Sequence 11, Appl
21	106	100.0	355	15	US-09-175-683B-9	Sequence 9, Appl
22	106	100.0	361	15	US-09-175-683-12	Sequence 12, Appl
23	106	100.0	361	15	US-09-175-683B-10	Sequence 10, Appl
24	106	100.0	376	15	US-09-175-683-13	Sequence 13, Appl
25	106	100.0	376	17	US-09-311-817-2	Sequence 2, Appl
26	106	100.0	379	15	US-09-175-683B-11	Sequence 11, Appl
27	106	100.0	384	19	US-09-500-376-8	Sequence 8, Appl
28	106	100.0	394	3	US-07-867-768A-2	Sequence 2, Appl
29	106	100.0	394	5	US-08-195-705-2	Sequence 2, Appl
30	106	100.0	394	19	US-09-500-376-2	Sequence 2, Appl
31	106	100.0	402	19	US-09-500-376-16	Sequence 16, Appl
32	106	100.0	594	9	US-08-593-006-2	Sequence 2, Appl
33	106	100.0	594	15	US-09-117-415-2	Sequence 2, Appl
34	106	100.0	613	9	US-08-593-006-22	Sequence 22, Appl
35	106	100.0	613	15	US-09-117-415-22	Sequence 22, Appl
36	106	100.0	631	9	US-08-593-006-18	Sequence 18, Appl
37	106	100.0	631	9	US-08-593-006-20	Sequence 20, Appl
38	106	100.0	631	15	US-09-117-415-18	Sequence 18, Appl
39	106	100.0	631	15	US-09-117-415-20	Sequence 20, Appl
40	106	100.0	649	9	US-08-593-006-16	Sequence 16, Appl
41	106	100.0	649	15	US-09-117-415-16	Sequence 16, Appl
42	106	100.0	1639	16	US-09-269-874-3	Sequence 3, Appl
43	103	97.2	375	3	US-07-867-768A-3	Sequence 3, Appl
44	103	97.2	375	5	US-08-195-705-3	Sequence 3, Appl
45	103	97.2	375	19	US-09-500-376-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-09-763-397A-12  
; Sequence 12, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasmain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; PRIOR FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-12

Query Match 100.0%; Score 106; DB 21; Length 19;  
Best Local Similarity 100.0%; Pred. No. 4.3e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EDGSGNGKKTCTKTPDS 19





Query Match 100.0%; Score 106; DB 15; Length 95;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKTCTCKPDS 19  
|||||  
DB 67 EDGSGNGKKTCTCKPDS 85

## RESULT 6

US-09-311-817-1  
; Sequence 1, Application US/09311817  
; GENERAL INFORMATION:  
; APPLICANT: Medical Research Council  
; TITLE OF INVENTION: Malaria Vaccine  
; FILE REFERENCE: 18396-1210  
; CURRENT APPLICATION NUMBER: US/09/311,817  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: GB 99099072.2  
; PRIOR FILING DATE: 1999-04-20  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-311-817-1

Query Match 100.0%; Score 106; DB 17; Length 96;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKTCTCKPDS 19  
|||||  
DB 65 EDGSGNGKKTCTCKPDS 83

## RESULT 7

US-09-125-031-10  
; Sequence 10, Application US/09125031  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: NATO, FARIDABANO  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031  
; PRIOR FILING DATE: 1999-03-10  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-125-031-10

Query Match 100.0%; Score 106; DB 15; Length 108;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKTCTCKPDS 19  
|||||  
DB 80 EDGSGNGKKTCTCKPDS 98

## RESULT 8

US-09-125-031A-10  
; Sequence 10, Application US/09125031A  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031A  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: PCT/FR97/00290  
; PRIOR FILING DATE: 1997-02-14  
; PRIOR APPLICATION NUMBER: FR96/01822  
; PRIOR FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-125-031A-10

Query Match 100.0%; Score 106; DB 15; Length 108;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKTCTCKPDS 19  
|||||  
DB 80 EDGSGNGKKTCTCKPDS 98

## RESULT 9

US-09-125-031B-10  
; Sequence 10, Application US/09125031B  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031B  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: PCT/FR97/00290  
; PRIOR FILING DATE: 1997-02-14  
; PRIOR APPLICATION NUMBER: FR96/01822  
; PRIOR FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-125-031B-10

Query Match 100.0%; Score 106; DB 15; Length 108;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKTCTCKPDS 19  
|||||  
DB 80 EDGSGNGKKTCTCKPDS 98

RESULT 10  
US-09-134-333-10  
; Sequence 10, Application US/09134333  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: NATO, FARIDABANO  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; TITLE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0135-0XCIP  
; CURRENT APPLICATION NUMBER: US/09/134,333  
; CURRENT FILING DATE: 1999-04-18  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-134-333-10

Query Match 100.0%; Score 106; DB 15; Length 108;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EDGSGNGKKTCTCKPDS 19  
DB 80 EDGSGNGKKTCTCKPDS 98  
|||||

RESULT 11  
US-09-125-031-5  
; Sequence 5, Application US/09125031  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: NATO, FARIDABANO  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; TITLE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031  
; CURRENT FILING DATE: 1999-03-10  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC  
US-09-125-031-5

Query Match 100.0%; Score 106; DB 15; Length 116;  
Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EDGSGNGKKTCTCKPDS 19  
DB 67 EDGSGNGKKTCTCKPDS 85  
|||||

RESULT 12  
US-09-125-031A-5  
; Sequence 5, Application US/09125031A  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; TITLE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031A  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: PCT/FR97/00290  
; PRIOR FILING DATE: 1997-02-14  
; PRIOR APPLICATION NUMBER: FR96/01822  
; PRIOR FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC  
US-09-125-031A-5

Query Match 100.0%; Score 106; DB 15; Length 116;  
Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EDGSGNGKKTCTCKPDS 19  
DB 67 EDGSGNGKKTCTCKPDS 85  
|||||

RESULT 13  
US-09-125-031B-5  
; Sequence 5, Application US/09125031B  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; TITLE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031B  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: PCT/FR97/00290  
; PRIOR FILING DATE: 1997-02-14  
; PRIOR APPLICATION NUMBER: FR96/01822  
; PRIOR FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC  
US-09-125-031B-5

Query Match 100.0%; Score 106; DB 15; Length 116;  
Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EDGSGNGKKTCTCKPDS 19  
DB 67 EDGSGNGKKTCTCKPDS 85  
|||||

QY 1 EDGSGNGKKITCCTKTPDS 19  
|||||  
Db 67 EDGSGNGKKITCCTKTPDS 85

RESULT 14  
US-09-134-333-5  
; Sequence 5, Application US/09134333  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: NATO, FARIDABANO  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; TITLE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0135-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/134,333  
; CURRENT FILING DATE: 1999-04-18  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC  
US-09-134-333-5

Query Match 100.0%; Score 106; DB 15; Length 116;  
Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCCTKTPDS 19  
|||||  
Db 67 EDGSGNGKKITCCTKTPDS 85

RESULT 15  
US-09-125-031-8  
; Sequence 8, Application US/09125031  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: NATO, FARIDABANO  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; TITLE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031  
; CURRENT FILING DATE: 1999-03-10  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-125-031-8

Query Match 100.0%; Score 106; DB 15; Length 127;  
Best Local Similarity 100.0%; Pred. No. 2.8e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKKITCCTKTPDS 19  
|||||  
Db 99 EDGSGNGKKITCCTKTPDS 117  
  
Search completed: January 29, 2002, 10:56:05  
Job time: 2401 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:12 ; Search time 120.95 Seconds  
(without alignments)  
5.749 Million cell updates/sec

Title: US-09-763-397A-12  
Perfect score: 106  
Sequence: 1 EDGSGNGKKTCECTKPD 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pcp.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pcp.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pcp.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pcp.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pcp.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pcp.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	100.0	96	5	US-09-978-756-1
2	106	100.0	376	5	US-09-978-756-2
3	106	100.0	594	5	US-09-117-415B-2
4	106	100.0	613	5	US-09-117-415B-22
5	106	100.0	631	5	US-09-117-415B-18
6	106	100.0	631	5	US-09-117-415B-20
7	106	100.0	649	5	US-09-117-415B-16
8	106	100.0	1602	5	US-09-269-874A-7
9	106	100.0	1621	5	US-09-269-874A-5
10	106	100.0	1639	5	US-09-269-874A-3
11	93	87.7	394	5	US-09-978-756-3
12	45	42.5	162	5	US-09-708-427-58009
13	43	40.6	609	5	US-09-801-368-50
14	43	40.6	609	5	US-09-487-558-50
15	42	39.6	230	5	US-09-708-427-31899
16	42	39.6	236	5	US-09-576-424-10
17	42	39.6	253	5	US-09-708-427-31898
18	42	39.6	345	5	US-09-708-427-8605
19	42	39.6	371	5	US-09-708-427-8604
20	42	39.6	660	5	US-09-708-427-13705
21	42	39.6	670	5	US-09-708-427-13704
22	42	39.6	738	5	US-09-708-427-13703
23	42	39.6	992	5	US-09-708-427-13240
24	42	39.6	1008	5	US-09-708-427-13239
25	42	39.6	1013	5	US-09-708-427-13238
26	41.5	39.2	155	5	US-09-708-427-13816

27	41.5	39.2	160	5	US-09-708-427-13815
28	41.5	39.2	263	5	US-09-708-427-13814
29	41	38.7	108	5	US-09-848-798A-68
30	41	38.7	108	5	US-09-848-798-68
31	41	38.7	114	5	US-09-974-449-8
32	41	38.7	194	5	US-09-708-427-72724
33	41	38.7	219	5	US-09-974-449-38
34	41	38.7	321	5	US-09-708-427-27613
35	41	38.7	463	6	US-10-005-368-57
36	41	38.7	1926	6	US-10-005-368-13
37	40.5	38.2	183	5	US-09-895-298A-70
38	40	37.7	110	5	US-09-848-798A-63
39	40	37.7	110	5	US-09-848-798-63
40	40	37.7	112	6	US-10-006-869-31
41	40	37.7	137	5	US-09-708-427-42719
42	40	37.7	154	5	US-09-708-427-42718
43	40	37.7	159	5	US-09-708-427-42717
44	40	37.7	221	6	US-10-001-857-202
45	40	37.7	263	5	US-09-708-427-34465

ALIGNMENTS

RESULT 1					
US-09-978-756-1					
Sequence 1, Application US/09978756					
GENERAL INFORMATION:					
APPLICANT: Holder, Anthony					
APPLICANT: Birdsall, Berry					
APPLICANT: Feeney, James					
APPLICANT: Morgan, William					
APPLICANT: Syed, Shabih					
TITLE OF INVENTION: Malaria Vaccine					
FILE REFERENCE: 18396/1005					
CURRENT APPLICATION NUMBER: US/09/978,756					
CURRENT FILING DATE: 2001-10-16					
PRIOR APPLICATION NUMBER: PCT/GB00/01558					
PRIOR FILING DATE: 2000-04-20					
PRIOR APPLICATION NUMBER: 09/311,817					
PRIOR FILING DATE: 1999-05-13					
PRIOR APPLICATION NUMBER: 2,271,451					
PRIOR FILING DATE: 1999-05-25					
PRIOR APPLICATION NUMBER: 9909072.2					
PRIOR FILING DATE: 1999-04-20					
NUMBER OF SEQ ID NOS: 3					
SOFTWARE: PatentIn version 3.0					
SEQ ID NO 1					
LENGTH: 96					
TYPE: PRT					
ORGANISM: Plasmodium falciparum					
US-09-978-756-1					
Query Match					
Best Local Similarity	100.0%	Score 106;	DB 5;	Length 96;	
Matches	19;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1	EDSGSNGKKTCECTKPD 19			
Db	65	EDSGSNGKKTCECTKPD 83			
RESULT 2					
US-09-978-756-2					
Sequence 2, Application US/09978756					
GENERAL INFORMATION:					
APPLICANT: Holder, Anthony					
APPLICANT: Birdsall, Berry					
APPLICANT: Feeney, James					
APPLICANT: Morgan, William					
APPLICANT: Syed, Shabih					
TITLE OF INVENTION: Malaria Vaccine					

```

FILE REFERENCE: 18396/1005
CURRENT APPLICATION NUMBER: US/09/978,756
CURRENT FILING DATE: 2001-10-16
PRIORITY APPLICATION NUMBER: PCT/GB00/01558
PRIORITY FILING DATE: 2000-04-20
PRIORITY APPLICATION NUMBER: 09/311,817
PRIORITY FILING DATE: 1999-05-13
PRIORITY APPLICATION NUMBER: 2,271,451
PRIORITY FILING DATE: 1999-05-25
PRIORITY APPLICATION NUMBER: 9909072.2
PRIORITY FILING DATE: 1999-04-20
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 376
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-09-978-756-2

Query Match      100.0%; Score 106; DB 5; Length 376;
Best Local Similarity 100.0%; Pred. No. 8.5e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EDGSGNGKKITCECTKPD 19
        |||||
Db      327 EDGSGNGKKITCECTKPD 345

RESULT 3
US-09-117-415B-2
; Sequence 2, Application US/091117415B
; GENERAL INFORMATION:
; APPLICANT: Davidson, Eugene
;               Yang, Shutong
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition
;                   of a MSAL Peptide
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESS: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC
; STREET: 714 Colorado Avenue
; CITY: Bridgeport
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPad (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/117,415B
; FILING DATE: 29-Jul-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coleman, Henry D.
; REGISTRATION NUMBER: 32,559
; REFERENCE/DOCKET NUMBER: R12-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 679-0090
; TELEFAX: (212) 679-9121
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 594 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-117-415B-2

Query Match      100.0%; Score 106; DB 5; Length 594;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;

```

```
;
;
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPad (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/117,415B
; FILING DATE: 29-Jul-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coleman, Henry D.
; REGISTRATION NUMBER: 32,559
; REFERENCE/DOCKET NUMBER: R12-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 679-0090
; TELEFAX: (212) 679-9121
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-117-415B-18

Query Match 100.0%; Score 106; DB 5; Length 631;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKITCCTKPD 19
   |||||
Db 600 EDGSGNGKITCCTKPD 618

RESULT 6
US-09-117-415B-20
; Sequence 20, Application US/09117415B
; GENERAL INFORMATION:
; APPLICANT: Davidson, Eugene
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition
; of a MSA1 Peptide
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC
; STREET: 714 Colorado Avenue
; CITY: Bridgeport
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPad (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/117,415B
; FILING DATE: 29-Jul-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coleman, Henry D.
; REGISTRATION NUMBER: 32,559
; REFERENCE/DOCKET NUMBER: R12-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 679-0090
; TELEFAX: (212) 679-9121
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-117-415B-16

Query Match 100.0%; Score 106; DB 5; Length 649;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDGSGNGKITCCTKPD 19
   |||||
Db 600 EDGSGNGKITCCTKPD 618

RESULT 8
US-09-269-874A-7
; Sequence 7, Application US/09269874A
; GENERAL INFORMATION:
; APPLICANT: Butard, Hermann
; TITLE OF INVENTION: Recombinant Process for Preparing a
; Complete Malaria Antigen, GP190/MSPI
; FILE REFERENCE: GRUE-003
```

```
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
```



; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..162  
; OTHER INFORMATION: Ceres Seq. ID 1941076  
US-09-708-427-58009

Query Match 42.5%; Score 45; DB 5; Length 162;  
Best Local Similarity 50.0%; Pred. No. 7.5;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 GSNKKITCECTKP 17  
|||: |||  
Db 128 GSGRRSPCTCTPP 141

RESULT 13  
US-09-801-368-50  
; Sequence 50, Application US/09801368  
; GENERAL INFORMATION:  
; APPLICANT: Busby, Robert  
; APPLICANT: Cali, Brian  
; APPLICANT: Hecht, Peter  
; APPLICANT: Holtzman, Doug  
; APPLICANT: Madden, Kevin  
; APPLICANT: Maxon, Mary  
; APPLICANT: Milne, Todd  
; APPLICANT: Norman, Thea  
; APPLICANT: Royer, John  
; APPLICANT: Salama, Sofie  
; APPLICANT: Sherman, Amir  
; APPLICANT: Silva, Jeff  
; APPLICANT: Summers, Eric  
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
; FILE REFERENCE: 109272.147  
; CURRENT APPLICATION NUMBER: US/09/801.368  
; CURRENT FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 09/487,558  
; PRIOR FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US 60/160,587  
; PRIOR FILING DATE: 1999-10-20  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 50  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-801-368-50

Query Match 40.6%; Score 43; DB 5; Length 609;  
Best Local Similarity 57.9%; Pred. No. 51;  
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

Qy 1 EDGSGNKKITCECTKPS 19  
|||: |||  
Db 5 EDFGSGKKE---TSPDS 19

RESULT 14  
US-09-487-558-50  
; Sequence 50, Application US/09487558  
; GENERAL INFORMATION:  
; APPLICANT: Busby, Robert  
; APPLICANT: Cali, Brian  
; APPLICANT: Hecht, Peter  
; APPLICANT: Holtzman, Doug  
; APPLICANT: Madden, Kevin  
; APPLICANT: Maxon, Mary  
; APPLICANT: Milne, Todd  
; APPLICANT: Norman, Thea  
; APPLICANT: Royer, John  
; APPLICANT: Salama, Sofie

; APPLICANT: Sherman, Amir  
; APPLICANT: Silva, Jeff  
; APPLICANT: Summers, Eric  
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production  
; FILE REFERENCE: 109272.147  
; CURRENT APPLICATION NUMBER: US/09/487,558  
; CURRENT FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US/09/801,368  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 09/487,558  
; PRIOR FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US 60/160,587  
; PRIOR FILING DATE: 1999-10-20  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 50  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-487-558-50

Query Match 40.6%; Score 43; DB 5; Length 609;  
Best Local Similarity 57.9%; Pred. No. 51;  
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

Qy 1 EDGSGNKKITCECTKPS 19  
|||: |||  
Db 5 EDFGSGKKE---TSPDS 19

RESULT 15  
US-09-708-427-31899  
; Sequence 31899, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31899  
; LENGTH: 230  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: misc.feature  
; LOCATION: 1..230  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..230  
; OTHER INFORMATION: Ceres Seq. ID 1833059  
US-09-708-427-31899

Query Match 39.6%; Score 42; DB 5; Length 230;  
Best Local Similarity 61.5%; Pred. No. 30;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GSGNKKITCECT 15  
|||: |||  
Db 14 SNSNKKLTCAAT 26

Search completed: January 29, 2002, 10:58:13  
Job time: 2509 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:04 ; Search time 1760.55 seconds  
(without alignments)  
2.839 Million cell updates/sec

Title: US-09-763-397A-11

Perfect score: 105

Sequence: 1 NSGCFRHLDERECKLL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 27767034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

#### Database :

Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query %	Score	Match	Length	DB ID	Description
1	105	100.0	18	21	US-09-763-397A-11	Sequence 11, Appl
2	105	100.0	95	15	US-09-125-031-2	Sequence 2, Appl
3	105	100.0	95	15	US-09-125-031A-2	Sequence 2, Appl
4	105	100.0	95	15	US-09-125-031B-2	Sequence 2, Appl
5	105	100.0	95	15	US-09-134-333-2	Sequence 2, Appl
6	105	100.0	96	17	US-09-311-817-1	Sequence 1, Appl
7	105	100.0	108	15	US-09-125-031-10	Sequence 10, Appl
8	105	100.0	108	15	US-09-125-031A-10	Sequence 10, Appl
9	105	100.0	108	15	US-09-125-031B-10	Sequence 10, Appl

10	105	100.0	108	15	US-09-134-333-10	Sequence 10, Appl
11	105	100.0	116	15	US-09-125-031-5	Sequence 5, Appl
12	105	100.0	116	15	US-09-125-031A-5	Sequence 5, Appl
13	105	100.0	116	15	US-09-125-031B-5	Sequence 5, Appl
14	105	100.0	116	15	US-09-134-333-5	Sequence 5, Appl
15	105	100.0	127	15	US-09-125-031-8	Sequence 8, Appl
16	105	100.0	127	15	US-09-125-031B-8	Sequence 8, Appl
17	105	100.0	127	15	US-09-134-333-8	Sequence 8, Appl
18	105	100.0	127	15	US-09-134-333-8	Sequence 8, Appl
19	105	100.0	350	21	US-09-763-397A-2	Sequence 2, Appl
20	105	100.0	355	15	US-09-175-683-11	Sequence 11, Appl
21	105	100.0	355	15	US-09-175-683B-9	Sequence 9, Appl
22	105	100.0	361	15	US-09-175-683-12	Sequence 12, Appl
23	105	100.0	361	15	US-09-175-683B-10	Sequence 10, Appl
24	105	100.0	375	3	US-07-867-768A-3	Sequence 3, Appl
25	105	100.0	375	5	US-08-195-705-3	Sequence 3, Appl
26	105	100.0	375	19	US-09-500-376-4	Sequence 4, Appl
27	105	100.0	375	21	US-09-710-000-8	Sequence 8, Appl
28	105	100.0	376	15	US-09-175-683-13	Sequence 13, Appl
29	105	100.0	376	17	US-09-311-817-2	Sequence 2, Appl
30	105	100.0	379	15	US-09-175-683B-11	Sequence 11, Appl
31	105	100.0	384	19	US-09-500-376-8	Sequence 8, Appl
32	105	100.0	394	3	US-07-867-768A-2	Sequence 2, Appl
33	105	100.0	394	3	US-07-867-768A-4	Sequence 4, Appl
34	105	100.0	394	5	US-08-195-705-2	Sequence 2, Appl
35	105	100.0	394	5	US-08-195-705-4	Sequence 4, Appl
36	105	100.0	394	19	US-09-500-376-2	Sequence 2, Appl
37	105	100.0	394	19	US-09-500-376-3	Sequence 3, Appl
38	105	100.0	402	19	US-09-500-376-16	Sequence 16, Appl
39	105	100.0	594	9	US-08-593-006-2	Sequence 2, Appl
40	105	100.0	594	15	US-09-117-415-2	Sequence 2, Appl
41	105	100.0	613	9	US-08-593-006-22	Sequence 22, Appl
42	105	100.0	613	15	US-09-117-415-22	Sequence 22, Appl
43	105	100.0	631	9	US-08-593-006-18	Sequence 18, Appl
44	105	100.0	631	9	US-08-593-006-20	Sequence 20, Appl
45	105	100.0	631	15	US-09-117-415-18	Sequence 18, Appl

#### ALIGNMENTS

RESULT 1  
US-09-763-397A-11  
; Sequence 11, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-11

Query Match 100.0% Score 105; DB 21; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.1e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSGCFRHLDERECKLL 18

Mon Feb 4 15:23:34 2002

us-09-763-397a-11.rapm

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18  
| | | | | | | | | | | | | | | | | |  
Db 17 NSGCFRHLDERECKCLL 34

RESULT 4  
US-09-125-031B-2  
; Sequence 2, Application US/09125031B  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; TITLE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125-031B  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: PCT/FR97/00290  
; PRIOR FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-125-031B-2

Query Match 100.0%; Score 105; DB 15; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18  
| | | | | | | | | | | | | | | | | |  
Db 17 NSGCFRHLDERECKCLL 34

RESULT 5  
US-09-134-333-2  
; Sequence 2, Application US/09134333  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; TITLE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0135-0XCIP  
; CURRENT APPLICATION NUMBER: US/09/134,333  
; CURRENT FILING DATE: 1999-04-18  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-134-333-2

Db 1 NSGCFRHLDERECKCLL 18  
| | | | | | | | | | | | | | | | | |

RESULT 2  
US-09-125-031-2  
; Sequence 2, Application US/09125031  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; TITLE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031  
; CURRENT FILING DATE: 1999-03-10  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-125-031-2

Query Match 100.0%; Score 105; DB 15; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18  
| | | | | | | | | | | | | | | | | |  
Db 17 NSGCFRHLDERECKCLL 34

RESULT 3  
US-09-125-031A-2  
; Sequence 2, Application US/09125031A  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; TITLE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031A  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: PCT/FR97/00290  
; PRIOR FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-125-031A-2

Query Match 100.0%; Score 105; DB 15; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;

Query Match 100.0%; Score 105; DB 15; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18  
|||||  
DB 17 NSGCFRHLDERECKLL 34

## RESULT 6

US-09-311-817-1  
; Sequence 1, Application US/09311817  
; GENERAL INFORMATION:  
; APPLICANT: Medical Research Council  
; TITLE OF INVENTION: Malaria Vaccine  
; FILE REFERENCE: 18396-1210  
; CURRENT APPLICATION NUMBER: US/09/311.817  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: GB 99099072.2  
; PRIOR FILING DATE: 1999-04-20  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-311-817-1

Query Match 100.0%; Score 105; DB 17; Length 96;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18  
|||||  
DB 15 NSGCFRHLDERECKLL 32

## RESULT 7

US-09-125-031-10  
; Sequence 10, Application US/09125031  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: NATO, FARIDABANO  
; APPLICANT: MENDIS, KAMINI  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031  
; CURRENT FILING DATE: 1999-03-10  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-125-031-10

Query Match 100.0%; Score 105; DB 15; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.3e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18  
|||||  
DB 30 NSGCFRHLDERECKLL 47

## RESULT 8

US-09-125-031A-10  
; Sequence 10, Application US/09125031A  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031A  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: PCT/FR97/00290  
; PRIOR FILING DATE: 1997-02-14  
; PRIOR APPLICATION NUMBER: FR96/01822  
; PRIOR FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-125-031A-10

Query Match 100.0%; Score 105; DB 15; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.3e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18  
|||||  
DB 30 NSGCFRHLDERECKLL 47

## RESULT 9

US-09-125-031B-10  
; Sequence 10, Application US/09125031B  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031B  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: PCT/FR97/00290  
; PRIOR FILING DATE: 1997-02-14  
; PRIOR APPLICATION NUMBER: FR96/01822  
; PRIOR FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-125-031B-10

Query Match 100.0%; Score 105; DB 15; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.3e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18  
|||||  
DB 30 NSGCFRHLDERECKLL 47

RESULT 10  
US-09-134-333-10  
; Sequence 10, Application US/09134333  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: NATO, FARIDABANO  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE REFERENCE: 0660-0135-0XCIP  
; CURRENT APPLICATION NUMBER: US/09/134,333  
; CURRENT FILING DATE: 1999-04-18  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-134-333-10

Query Match 100.0%; Score 105; DB 15; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.3e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18  
Db 30 NSGCFRHLDERECKCLL 47

RESULT 11  
US-09-125-031-5  
; Sequence 5, Application US/09125031  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: NATO, FARIDABANO  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031  
; CURRENT FILING DATE: 1999-03-10  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC  
US-09-125-031-5

Query Match 100.0%; Score 105; DB 15; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18  
Db 17 NSGCFRHLDERECKCLL 34

RESULT 12  
US-09-125-031A-5  
; Sequence 5, Application US/09125031A  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031A  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: PCT/FR97/00290  
; PRIOR FILING DATE: 1997-02-14  
; PRIOR APPLICATION NUMBER: FR96/01822  
; PRIOR FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC  
US-09-125-031A-5

Query Match 100.0%; Score 105; DB 15; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKCLL 18  
Db 17 NSGCFRHLDERECKCLL 34

RESULT 13  
US-09-125-031B-5  
; Sequence 5, Application US/09125031B  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; APPLICANT: NATO, FARIDABANO  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031B  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: PCT/FR97/00290  
; PRIOR FILING DATE: 1997-02-14  
; PRIOR APPLICATION NUMBER: FR96/01822  
; PRIOR FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC  
US-09-125-031B-5

Query Match 100.0%; Score 105; DB 15; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18  
          |||||  
Db 17 NSGCFRHLDERECKLL 34

RESULT 14  
US-09-134-333-5  
; Sequence 5, Application US/09134333  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: NATO, FARIDABANO  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0135-0XCIP  
; CURRENT APPLICATION NUMBER: US/09/134,333  
; CURRENT FILING DATE: 1999-04-18  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC  
US-09-134-333-5

Query Match 100.0%; Score 105; DB 15; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18  
          |||||  
Db 17 NSGCFRHLDERECKLL 34

RESULT 15  
US-09-125-031-8  
; Sequence 8, Application US/09125031  
; GENERAL INFORMATION:  
; APPLICANT: LONGACRE-ANDRE, SHIRLEY  
; APPLICANT: ROTH, CHARLES  
; APPLICANT: NATO, FARIDABANO  
; APPLICANT: BARNWELL, JOHN  
; APPLICANT: MENDIS, KAMINI  
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF  
; FILE OF INVENTION: PLASMODIUM MSP-1  
; FILE REFERENCE: 0660-0139-0XPCT  
; CURRENT APPLICATION NUMBER: US/09/125,031  
; CURRENT FILING DATE: 1999-03-10  
; EARLIER APPLICATION NUMBER: PCT/FR97/00290  
; EARLIER FILING DATE: 1997-02-14  
; EARLIER APPLICATION NUMBER: FR96/01822  
; EARLIER FILING DATE: 1996-02-14  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-125-031-8

Query Match 100.0%; Score 105; DB 15; Length 127;  
Best Local Similarity 100.0%; Pred. No. 1.5e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18  
          |||||  
Db 49 NSGCFRHLDERECKLL 66

Search completed: January 29, 2002, 10:56:05  
Job time: 2401 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:12 ; Search time 120.95 Seconds  
(without alignments)  
5.446 Million cell updates/sec

Title: US-09-763-397A-11

Perfect score: 105  
Sequence: 1 NSGCFRHLDERECKLL 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	105	100.0	96	5	US-09-978-756-1
2	105	100.0	376	5	US-09-978-756-2
3	105	100.0	394	5	US-09-978-756-3
4	105	100.0	594	5	US-09-117-4158-2
5	105	100.0	613	5	US-09-117-4158-22
6	105	100.0	631	5	US-09-117-4158-18
7	105	100.0	631	5	US-09-117-4158-20
8	105	100.0	649	5	US-09-117-4158-16
9	105	100.0	1602	5	US-09-269-874A-7
10	105	100.0	1621	5	US-09-269-874A-5
11	105	100.0	1639	5	US-09-269-874A-3
12	45	42.9	85	5	US-09-605-703B-1594
13	44	41.9	1315	5	US-09-708-427-32716
14	44	41.9	1342	5	US-09-708-427-32715
15	44	41.9	1444	5	US-09-708-427-32714
16	42	40.0	876	5	US-09-897-516-5022
17	41	39.0	185	5	US-09-708-427-63903
18	41	39.0	185	5	US-09-708-427-82641
19	41	39.0	234	5	US-09-708-427-63902
20	41	39.0	234	5	US-09-708-427-82640
21	41	39.0	550	5	US-09-605-703B-1906
22	41	39.0	588	5	US-09-908-193-34
23	41	39.0	754	5	US-09-908-193-32
24	41	39.0	776	5	US-09-908-193-33
25	41	39.0	778	5	US-09-908-193-8
26	41	39.0	779	5	US-09-908-193-10

27	40.5	38.6	578	5	US-09-783-931-13	Sequence 13, Appl
28	40	38.1	108	5	US-09-708-427-39610	Sequence 39610, A
29	40	38.1	398	5	US-09-708-427-863	Sequence 863, App
30	40	38.1	451	5	US-09-708-427-862	Sequence 862, App
31	39.5	37.6	886	5	US-09-992-647-11	Sequence 11, Appl
32	39	37.1	108	5	US-09-708-427-49464	Sequence 49464, A
33	39	37.1	114	5	US-09-708-427-49463	Sequence 49463, A
34	39	37.1	136	5	US-09-620-394B-936	Sequence 936, App
35	39	37.1	136	5	US-09-708-427-77213	Sequence 77213, A
36	39	37.1	136	5	US-09-708-427-79661	Sequence 79661, A
37	39	37.1	142	5	US-09-620-394B-935	Sequence 935, App
38	39	37.1	142	5	US-09-708-427-77212	Sequence 77212, A
39	39	37.1	142	5	US-09-708-427-79660	Sequence 79660, A
40	39	37.1	163	5	US-09-708-427-49462	Sequence 49462, A
41	39	37.1	168	5	US-09-620-394B-934	Sequence 934, App
42	39	37.1	303	5	US-09-620-394B-7490	Sequence 7490, Ap
43	39	37.1	323	5	US-09-620-394B-7489	Sequence 7489, Ap
44	39	37.1	435	5	US-09-708-427-17436	Sequence 17436, A
45	39	37.1	513	5	US-09-708-427-17435	Sequence 17435, A

ALIGNMENTS

RESULT 1

US-09-978-756-1  
; Sequence 1, Application US/09978756  
; GENERAL INFORMATION:  
; APPLICANT: Holder, Anthony  
; APPLICANT: Birdsall, Berry  
; APPLICANT: Feeney, James  
; APPLICANT: Morgan, William  
; APPLICANT: Syed, Shabih  
; TITLE OF INVENTION: Malaria Vaccine  
; FILE REFERENCE: 18396/1005  
; CURRENT APPLICATION NUMBER: US/09/978,756  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: PCT/GB00/01558  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 09/311,817  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: 2,271,451  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 9909072.2  
; PRIOR FILING DATE: 1999-04-20  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-978-756-1

Query Match 100.0%; Score 105; DB 5; Length 96;  
Best Local Similarity 100.0%; Pred. No. 5.5e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18  
Db 15 NSGCFRHLDERECKLL 32

RESULT 2

US-09-978-756-2  
; Sequence 2, Application US/09978756  
; GENERAL INFORMATION:  
; APPLICANT: Holder, Anthony  
; APPLICANT: Birdsall, Berry  
; APPLICANT: Feeney, James  
; APPLICANT: Morgan, William  
; APPLICANT: Syed, Shabih  
; TITLE OF INVENTION: Malaria Vaccine

Mon Feb 4 15:23:34 2002

```
FILE REFERENCE: 18396/1005
CURRENT APPLICATION NUMBER: US/09/978,756
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: PCT/GB00/01558
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: 09/311,817
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: 2,271,451
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 9909072.2
PRIOR FILING DATE: 1999-04-20
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 376
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-09-978-756-2

Query Match      100.0%; Score 105; DB 5; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 277 NSGCFRHLDERECKLL 294

RESULT 3
US-09-978-756-3
Sequence 3, Application US/09978756
GENERAL INFORMATION:
APPLICANT: Holder, Anthony
APPLICANT: Birdsell, Berry
APPLICANT: Peeney, James
APPLICANT: Morgan, William
APPLICANT: Syed, Shabih
TITLE OF INVENTION: Malaria Vaccine
FILE REFERENCE: 18396/1005
CURRENT APPLICATION NUMBER: US/09/978,756
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: PCT/GB00/01558
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: 09/311,817
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: 2,271,451
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 9909072.2
PRIOR FILING DATE: 1999-04-20
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 394
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-09-978-756-3

Query Match      100.0%; Score 105; DB 5; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 295 NSGCFRHLDERECKLL 312

RESULT 4
US-09-117-415B-2
Sequence 2, Application US/09117415B
GENERAL INFORMATION:
APPLICANT: Davidson, Eugene
```

```
Yang, Shutong
TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition
of a MSAl Peptide
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC
STREET: 714 Colorado Avenue
CITY: Bridgeport
STATE: Connecticut
COUNTRY: USA
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPad (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/117,415B
FILING DATE: 29-Jul-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coleman, Henry D.
REGISTRATION NUMBER: 32,559
REFERENCE/DOCKET NUMBER: R12-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 679-0090
TELEFAX: (212) 679-9121
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 594 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-117-415B-2

Query Match      100.0%; Score 105; DB 5; Length 594;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSGCFRHLDERECKLL 18
DB 495 NSGCFRHLDERECKLL 512

RESULT 5
US-09-117-415B-22
Sequence 22, Application US/09117415B
GENERAL INFORMATION:
APPLICANT: Davidson, Eugene
Yang, Shutong
TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition
of a MSAl Peptide
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC
STREET: 714 Colorado Avenue
CITY: Bridgeport
STATE: Connecticut
COUNTRY: USA
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPad (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/117,415B
FILING DATE: 29-Jul-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coleman, Henry D.
```

REGISTRATION NUMBER: 32,559  
REFERENCE/DOCKET NUMBER: R12-030  
TELEPHONE: (212) 679-0090  
TELEFAX: (212) 679-9121  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 613 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-117-415B-22

Query Match 100.0%; Score 105; DB 5; Length 613;  
Best Local Similarity 100.0%; Pred. No. 2.7e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18  
Db 532 NSGCFRHLDERECKLL 549

## RESULT 6

US-09-117-415B-18  
; Sequence 18, Application US/09117415B  
; GENERAL INFORMATION:  
; APPLICANT: Davidson, Eugene  
; Yang, Shutong  
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
; of a MSAL Peptide  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
; STREET: 714 Colorado Avenue  
; CITY: Bridgeport  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 10017

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPad (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/117,415B  
FILING DATE: 29-Jul-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coleman, Henry D.  
REGISTRATION NUMBER: 32,559  
REFERENCE/DOCKET NUMBER: R12-030  
TELEPHONE: (212) 679-0090  
TELEFAX: (212) 679-9121  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 631 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-117-415B-18

Query Match 100.0%; Score 105; DB 5; Length 631;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18  
Db 550 NSGCFRHLDERECKLL 567

## RESULT 7

US-09-117-415B-20  
; Sequence 20, Application US/09117415B  
; GENERAL INFORMATION:  
; APPLICANT: Davidson, Eugene  
; Yang, Shutong  
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
; of a MSAL Peptide  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
; STREET: 714 Colorado Avenue  
; CITY: Bridgeport  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 10017

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPad (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/117,415B  
FILING DATE: 29-Jul-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coleman, Henry D.  
REGISTRATION NUMBER: 32,559  
REFERENCE/DOCKET NUMBER: R12-030  
TELEPHONE: (212) 679-0090  
TELEFAX: (212) 679-9121  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 631 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-117-415B-20

Query Match 100.0%; Score 105; DB 5; Length 631;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSGCFRHLDERECKLL 18  
Db 532 NSGCFRHLDERECKLL 549

## RESULT 8

US-09-117-415B-16  
; Sequence 16, Application US/09117415B  
; GENERAL INFORMATION:  
; APPLICANT: Davidson, Eugene  
; Yang, Shutong  
; TITLE OF INVENTION: Malaria Vaccine Based Upon the Addition  
; of a MSAL Peptide  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Henry D. Coleman, COLEMAN SUDOL SAPONE, PC  
; STREET: 714 Colorado Avenue  
; CITY: Bridgeport  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 10017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

Mon Feb 4 15:23:34 2002

```

; SOFTWARE: WordPad (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/117,415B
; FILING DATE: 29-Jul-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coleman, Henry D.
; REGISTRATION NUMBER: 32,559
; REFERENCE/DOCKET NUMBER: R12-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 679-0090
; TELEFAX: (212) 679-9121
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-117-415B-16

```

```

Query Match 100.0%; Score 105; DB 5; Length 649;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSGCFRHLDERECKCLL 18
Db 550 NSGCFRHLDERECKCLL 567

```

```

RESULT 9
US-09-269-874A-7
; Sequence 7, Application US/09269874A
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; TITLE OF INVENTION: Recombinant Process for Preparing a
; TITLE OF INVENTION: Complete Malaria Antigen, GP190/MSPl
; FILE REFERENCE: GRUE-003
; CURRENT APPLICATION NUMBER: US/09/269,874A
; CURRENT FILING DATE: 1999-08-02
; PRIOR APPLICATION NUMBER: PCT/EP97/05441
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: DE 19640817.2
; PRIOR FILING DATE: 1996-10-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1602
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-269-874A-7

```

```

Query Match 100.0%; Score 105; DB 5; Length 1602;
Best Local Similarity 100.0%; Pred. No. 6.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSGCFRHLDERECKCLL 18
Db 1521 NSGCFRHLDERECKCLL 1538

```

```

RESULT 10
US-09-269-874A-5
; Sequence 5, Application US/09269874A.
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; TITLE OF INVENTION: Recombinant Process for Preparing a
; TITLE OF INVENTION: Complete Malaria Antigen, GP190/MSPl
; FILE REFERENCE: GRUE-003
; CURRENT APPLICATION NUMBER: US/09/269,874A
; CURRENT FILING DATE: 1999-08-02

```

```

; PRIOR APPLICATION NUMBER: PCT/EP97/05441
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: DE 19640817.2
; PRIOR FILING DATE: 1996-10-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1621
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-269-874A-5

```

```

Query Match 100.0%; Score 105; DB 5; Length 1621;
Best Local Similarity 100.0%; Pred. No. 6.2e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSGCFRHLDERECKCLL 18
Db 1540 NSGCFRHLDERECKCLL 1557

```

```

RESULT 11
US-09-269-874A-3
; Sequence 3, Application US/09269874A
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; TITLE OF INVENTION: Recombinant Process for Preparing a
; TITLE OF INVENTION: Complete Malaria Antigen, GP190/MSPl
; FILE REFERENCE: GRUE-003
; CURRENT APPLICATION NUMBER: US/09/269,874A
; CURRENT FILING DATE: 1999-08-02
; PRIOR APPLICATION NUMBER: PCT/EP97/05441
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: DE 19640817.2
; PRIOR FILING DATE: 1996-10-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1639
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-269-874A-3

```

```

Query Match 100.0%; Score 105; DB 5; Length 1639;
Best Local Similarity 100.0%; Pred. No. 6.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSGCFRHLDERECKCLL 18
Db 1540 NSGCFRHLDERECKCLL 1557

```

```

RESULT 12
US-09-605-703B-1594
; Sequence 1594, Application US/09605703B
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934

```

; SEQ ID NO 1594  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-605-703B-1594

Query Match 42.9%; Score 45; DB 5; Length 85;  
Best Local Similarity 47.1%; Pred. No. 4.3;  
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 SGCFRHLDERECKCLL 18  
: | | | | | :  
Db 23 TGVIRLDEGEDCKAVV 39

RESULT 13  
US-09-708-427-32716  
; Sequence 32716, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32716  
; LENGTH: 1315  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1315  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1315  
; OTHER INFORMATION: Ceres Seq. ID 1834860  
US-09-708-427-32716

Query Match 41.9%; Score 44; DB 5; Length 1315;  
Best Local Similarity 52.9%; Pred. No. 64;  
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 3 GCFRHL-----DERECK 15  
| | | | | :  
Db 928 GCFSHLKTVFQEELECR 944

RESULT 14  
US-09-708-427-32715  
; Sequence 32715, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32715  
; LENGTH: 1342  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1342  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1342

; OTHER INFORMATION: Ceres Seq. ID 1834859  
US-09-708-427-32715

Query Match 41.9%; Score 44; DB 5; Length 1342;  
Best Local Similarity 52.9%; Pred. No. 65;  
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 3 GCFRHL-----DERECK 15  
| | | | | :  
Db 955 GCFSHLKTVFQEELECR 971

RESULT 15  
US-09-708-427-32714  
; Sequence 32714, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32714  
; LENGTH: 1444  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1444  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..1444  
; OTHER INFORMATION: Ceres Seq. ID 1834858  
US-09-708-427-32714

Query Match 41.9%; Score 44; DB 5; Length 1444;  
Best Local Similarity 52.9%; Pred. No. 69;  
Matches 9; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 3 GCFRHL-----DERECK 15  
| | | | | :  
Db 1057 GCFSHLKTVFQEELECR 1073

Search completed: January 29, 2002, 10:58:12  
Job time: 2508 sec

us-09-763-397a-11.rapn

Mon Feb 4 15:23:34 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:03 ; Search time 1760.55 seconds  
(without alignments)  
1.262 Million cell updates/sec

Title: US-09-763-397A-10  
Perfect score: 44  
Sequence: 1 KPNDKSLY 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main : \*  
1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	8	5	US-08-197-484-19
2	44	100.0	8	7	US-08-318-856-2
3	44	100.0	8	12	US-08-820-360-19
4	44	100.0	8	18	US-09-454-204A-2
5	44	100.0	8	21	US-09-763-397A-10
6	44	100.0	229	18	US-09-454-204A-41
7	44	100.0	350	21	US-09-763-397A-2
8	36	81.8	67	22	US-09-834-366-22032
9	36	81.8	67	24	US-60-147-499-7128

10	36	81.8	67	24	US-60-197-873-22032	Sequence 22032, A
11	36	81.8	514	1	PCT-US01-03782A-314	Sequence 314, App
12	36	81.8	586	21	US-09-758-441-318	Sequence 318, App
13	36	81.8	650	24	US-60-324-631-3037	Sequence 3037, Ap
14	35	79.5	1100	21	US-09-733-089-19706	Sequence 19706, A
15	35	79.5	1100	22	US-09-816-660-19706	Sequence 19706, A
16	34	77.3	84	20	US-09-617-681A-332	Sequence 332, App
17	34	77.3	95	20	US-09-617-681A-330	Sequence 330, App
18	34	77.3	219	24	US-60-215-161-6362	Sequence 6362, Ap
19	34	77.3	238	20	US-09-688-051-1027	Sequence 1027, Ap
20	34	77.3	247	16	US-09-252-691-6281	Sequence 6281, Ap
21	34	77.3	247	16	US-09-252-691C-6281	Sequence 6281, Ap
22	34	77.3	260	20	US-09-688-051-1026	Sequence 1026, Ap
23	34	77.3	275	20	US-09-688-051-1025	Sequence 1025, Ap
24	34	77.3	380	1	PCT-US98-06371-1200	Sequence 1200, Ap
25	34	77.3	380	13	US-08-902-615A-466	Sequence 466, App
26	34	77.3	385	15	US-09-107-532-5168	Sequence 5168, App
27	34	77.3	385	15	US-09-107-532A-5168	Sequence 5168, Ap
28	34	77.3	418	18	US-09-489-039A-11946	Sequence 11946, A
29	34	77.3	7618	24	US-60-167-217-1269	Sequence 1269, Ap
30	34	77.3	7676	24	US-60-173-464-1020	Sequence 1020, Ap
31	33	75.0	290	24	US-60-167-217-415	Sequence 415, App
32	33	75.0	290	24	US-60-191-637-435	Sequence 435, App
33	33	75.0	296	16	US-09-270-767-43860	Sequence 43860, A
34	33	75.0	313	1	PCT-US99-22853B-2246	Sequence 2246, Ap
35	33	75.0	366	1	PCT-US99-22853B-2245	Sequence 2245, Ap
36	33	75.0	399	1	PCT-US99-22853B-2244	Sequence 2244, Ap
37	33	75.0	444	18	US-09-413-198-2464	Sequence 2464, Ap
38	33	75.0	647	16	US-09-248-796-26574	Sequence 26574, A
39	33	75.0	1089	23	US-09-902-540-12094	Sequence 12094, A
40	33	75.0	1355	24	US-60-243-468-727	Sequence 727, App
41	32	72.7	62	15	US-09-134-001C-3683	Sequence 3683, Ap
42	32	72.7	72	1	PCT-US01-18569-2797	Sequence 2797, Ap
43	32	72.7	96	1	PCT-US01-08631-4197	Sequence 4197, A
44	32	72.7	112	21	US-09-758-472-6778	Sequence 6778, Ap
45	32	72.7	129	1	PCT-US01-08631-44592	Sequence 44592, A

## ALIGNMENTS

RESULT 1  
US-08-197-484-19  
; Sequence 19, Application US/08197484  
; GENERAL INFORMATION:  
; APPLICANT: VITIELLO, Maria A.  
; APPLICANT: CHESTNUT, Robert W.  
; APPLICANT: SETTE, Alessandro D.  
; APPLICANT: CELIS, Esteban  
; TITLE OF INVENTION: GRAY, Howard  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; TITLE OF INVENTION: CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/197,484  
; FILING DATE: 16-FEB-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/935,811  
; FILING DATE: 26-AUG-1992  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,682  
FILING DATE: 29-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,568  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 14137-26-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (206) 623-6793  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-197-484-19

Query Match 100.0%; Score 44; DB 5; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
| | | | | | | |  
Db 1 KPNDKSLY 8

## RESULT 2

US-08-318-856-2  
Sequence 2, Application US/08318856  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Peptides of an Antigen, Capable of  
Recognition by or Induction of Cytotoxic T Lymphocytes, and  
TITLE OF INVENTION: Method of Identification  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1+  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,856  
FILING DATE: October 3, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 92 08 068.8  
FILING DATE: April 3, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 92 17 704.7  
FILING DATE: August 20, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/GB93/00711  
FILING DATE: April 5, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew Jacob  
REGISTRATION NUMBER: 25,154  
REFERENCE/DOCKET NUMBER: 263-PP1R1577US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)-371-8850

TELEFAX: (202) 371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium falciparum  
US-08-318-856-2

Query Match 100.0%; Score 44; DB 7; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
| | | | | | | |  
Db 1 KPNDKSLY 8

## RESULT 3

US-08-820-360-19  
Sequence 19, Application US/08820360  
GENERAL INFORMATION:  
APPLICANT: VITIELLO, Maria A.  
APPLICANT: CHESTNUT, Robert W.  
APPLICANT: SETTE, Alessandro D.  
APPLICANT: CELIS, Esteban  
APPLICANT: GRAY, Howard  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
TITLE OF INVENTION: CTL IMMUNITY  
NUMBER OF SEQUENCES: 154  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/820,360  
FILING DATE: 12-MAR-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/013,366  
FILING DATE: 13-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/197,484  
FILING DATE: 16-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/935,811  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,682  
FILING DATE: 29-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,568  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 14137-26-4



; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 467-9600  
; TELEFAX: (206) 623-6793  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-820-360-19

Query Match 100.0%; Score 44; DB 12; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
DB 1 KPNDKSLY 8

## RESULT 4

US-09-454-204A-2  
; Sequence 2, Application US/09454204A  
; GENERAL INFORMATION:  
; APPLICANT: McMichael, Andrew  
; APPLICANT: Hill, Adrian V.S.  
; APPLICANT: Gilbert, Sarah C.  
; APPLICANT: Schneider, Jorg  
; APPLICANT: Plebanski, Magdalena  
; APPLICANT: Hanke, Tomas  
; APPLICANT: Smith, Geoffrey L.  
; APPLICANT: Blanchard, Tom  
; TITLE OF INVENTION: Methods and Reagents for Vaccination  
; FILE OF INVENTION: Which Generate A CD8 T Cell Immune Response  
; FILE REFERENCE: 2907.1000-000  
; CURRENT FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: US/09/454,204A  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: PCT/GB98/01681  
; PRIOR FILING DATE: 1997-06-09  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: CTL Epitope of the Malaria String  
US-09-454-204A-2

Query Match 100.0%; Score 44; DB 18; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
DB 1 KPNDKSLY 8

## RESULT 5

US-09-763-397A-10  
; Sequence 10, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.

; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-10

Query Match 100.0%; Score 44; DB 21; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
DB 1 KPNDKSLY 8

## RESULT 6

US-09-454-204A-41  
; Sequence 41, Application US/09454204A  
; GENERAL INFORMATION:  
; APPLICANT: McMichael, Andrew  
; APPLICANT: Hill, Adrian V.S.  
; APPLICANT: Gilbert, Sarah C.  
; APPLICANT: Schneider, Jorg  
; APPLICANT: Plebanski, Magdalena  
; APPLICANT: Hanke, Tomas  
; APPLICANT: Smith, Geoffrey L.  
; APPLICANT: Blanchard, Tom  
; TITLE OF INVENTION: Methods and Reagents for Vaccination  
; FILE OF INVENTION: Which Generate A CD8 T Cell Immune Response  
; FILE REFERENCE: 2907.1000-000  
; CURRENT FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: US/09/454,204A  
; CURRENT FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: PCT/GB98/01681  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: GB 97 11957.2  
; PRIOR FILING DATE: 1997-06-09  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 229  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Complete Epitope of Malaria String  
US-09-454-204A-41

Query Match 100.0%; Score 44; DB 18; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
DB 29 KPNDKSLY 36

## RESULT 7

US-09-763-397A-2  
; Sequence 2, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for

```
; APPLICANT: Control and Prevention
; APPLICANT: Lal, Altaf A.
; APPLICANT: Ping Shi, Ya
; APPLICANT: Hasnain, Seyed E.
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci
; FILE REFERENCE: 6395-57049
; CURRENT APPLICATION NUMBER: US/09/763,397A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/097,703
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: PCT / US99/18869
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant DNA/Protein
US-09-763-397A-2

Query Match      100.0%; Score 44; DB 21; Length 350;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY      1 KPNDKSLY 8
Db      184 KPNDKSLY 191

RESULT 8
US-09-834-366-22032
; Sequence 22032, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 22032
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-366-22032

Query Match      81.8%; Score 36; DB 22; Length 67;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Caps 0;

QY      1 KPNDKSLY 8
Db      58 KPNDKLIY 65

RESULT 9
US-60-147-499-7128
; Sequence 7128, Application US/60147499
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
```

```
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/60/147,499
; CURRENT FILING DATE: 1999-08-05
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7128
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-147-499-7128

Query Match      81.8%; Score 36; DB 24; Length 67;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Caps 0;

QY      1 KPNDKSLY 8
Db      58 KPNDKLIY 65

RESULT 10
US-60-197-873-22032
; Sequence 22032, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 22032
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-197-873-22032

Query Match      81.8%; Score 36; DB 24; Length 67;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Caps 0;

QY      1 KPNDKSLY 8
Db      58 KPNDKLIY 65

RESULT 11
PCT-US01-03782A-314
; Sequence 314, Application PC/TUS0103782A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Ford, John E et al
; TITLE OF INVENTION: Novel Bone Marrow Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-040
; CURRENT APPLICATION NUMBER: PCT/US01/03782A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/250,583
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: Custom
; SEQ ID NO 314
```

; LENGTH: 514  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(514)  
; OTHER INFORMATION: Xaa = any amino acid or nothing  
PCT-US01-03782A-314

Query Match 81.8%; Score 36; DB 1; Length 514;  
Best Local Similarity 85.7%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8  
| | | | |  
DB 93 PNDKSLY 99

RESULT 12  
US-09-758-441-318  
; Sequence 318, Application US/09758441  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PM040  
; CURRENT APPLICATION NUMBER: US/09/758,441  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; NUMBER OF SEQ ID NOS: 328  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 318  
; LENGTH: 586

; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (9)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (11)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (50)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (115)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-758-441-318

Query Match 81.8%; Score 36; DB 21; Length 586;  
Best Local Similarity 85.7%; Pred. No. 2.5e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8  
| | | | |  
DB 29 PNDKSLY 35

RESULT 13  
US-60-324-631-3037  
; Sequence 3037, Application US/60324631  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Felyan  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Jian-Rui

; APPLICANT: Wehrman, Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ghosh, Malabika  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Ma, Yunding  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Wang, Gezhi  
; APPLICANT: Haley-Vicente, Dana  
; APPLICANT: Drmanac, Radoje T  
; TITLE OF INVENTION: Novel Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 810  
; CURRENT APPLICATION NUMBER: US/60/324,631  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: US 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: PCT/US00/35017  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: PCT/US01/02623  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: US 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: US 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: PCT/US01/03800  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: US 09/515,126  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: US 09/577,409  
; PRIOR FILING DATE: 2000-05-18  
; PRIOR APPLICATION NUMBER: PCT/US01/04927  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: US 09/519,705  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: US 09/574,454  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: PCT/US01/04941  
; PRIOR FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: US 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: US 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 09/552,929  
; PRIOR FILING DATE: 2000-04-18  
; PRIOR APPLICATION NUMBER: US 09/770,160  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: PCT/US01/08656  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: US 09/577,408  
; PRIOR FILING DATE: 2000-05-18  
; PRIOR APPLICATION NUMBER: PCT/US01/14827  
; PRIOR FILING DATE: 2001-05-16  
; NUMBER OF SEQ ID NOS: 3334  
; SOFTWARE: pt\_FL\_genes Version 6.0  
; SEQ ID NO 3037  
; LENGTH: 650  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-324-631-3037

Query Match 81.8%; Score 36; DB 24; Length 650;  
Best Local Similarity 85.7%; Pred. No. 2.9e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8

Db 93 PNDKSLY 99  
| | | : | | |

## RESULT 14

US-09-733-089-19706  
; Sequence 19706, Application US/09733089  
; GENERAL INFORMATION:  
; APPLICANT: Dotson, Stanton B.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Lutfiyya, Linda L.  
; APPLICANT: McIninch, James  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
; TITLE OF INVENTION: Transcription In Plants  
; FILE REFERENCE: 38-21(15300)D  
; CURRENT APPLICATION NUMBER: US/09/733,089  
; CURRENT FILING DATE: 2000-12-11  
; PRIOR APPLICATION NUMBER: US 09/474,435  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: US 09/654,617  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 09/620,392  
; PRIOR FILING DATE: 2000-07-19  
; NUMBER OF SEQ ID NOS: 24143  
; SEQ ID NO 19706  
; LENGTH: 1100  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-09-733-089-19706

Query Match 79.5%; Score 35; DB 21; Length 1100;  
Best Local Similarity 75.0%; Pred. No. 8.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KPNKSLY 8  
: | | | | |  
Db 638 RPTDKSLY 645

## RESULT 15

US-09-816-660-19706  
; Sequence 19706, Application US/09816660  
; GENERAL INFORMATION:  
; APPLICANT: Dotson, Stanton B.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Lutfiyya, Linda L.  
; APPLICANT: McIninch, James  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
; TITLE OF INVENTION: Transcription In Plants  
; FILE REFERENCE: 38-21(15300)D  
; CURRENT APPLICATION NUMBER: US/09/816,660  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: US 09/474,435  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: US 09/654,617  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 09/733,089  
; PRIOR FILING DATE: 2000-12-11  
; PRIOR APPLICATION NUMBER: )  
; PRIOR FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 09/620,392  
; PRIOR FILING DATE: 2000-07-19  
; NUMBER OF SEQ ID NOS: 24143  
; SEQ ID NO 19706  
; LENGTH: 1100  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-09-816-660-19706

Query Match 79.5%; Score 35; DB 22; Length 1100;  
Best Local Similarity 75.0%; Pred. No. 8.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 KPNKSLY 8  
: | | | | |  
Db 638 RPTDKSLY 645

Search completed: January 29, 2002, 10:56:04  
Job time: 2400 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:11 ; Search time 120.95 Seconds  
(without alignments)  
2.421 Million cell updates/sec

Title: US-09-763-397A-10  
Perfect score: 44  
Sequence: 1 KPNDKSLY 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	77.3	219	5	US-09-897-516-6362
2	34	77.3	380	5	US-09-815-242-11291
3	32	72.7	290	5	US-09-708-427-21872
4	32	72.7	300	5	US-09-897-516-5176
5	32	72.7	334	5	US-09-708-427-21871
6	32	72.7	341	5	US-09-708-427-21870
7	32	72.7	969	5	US-09-981-353-122
8	32	72.7	2115	7	US-60-337-444-4
9	32	72.7	2150	7	US-60-337-444-27
10	31	70.5	220	5	US-09-637-780B-196
11	31	70.5	220	5	US-09-708-427-72696
12	31	70.5	236	5	US-09-637-780B-195
13	31	70.5	236	5	US-09-708-427-72695
14	31	70.5	291	5	US-09-637-780B-194
15	31	70.5	292	5	US-09-708-427-72694
16	31	70.5	293	5	US-09-605-703B-54
17	31	70.5	319	5	US-09-605-703B-56
18	31	70.5	365	5	US-09-897-516-5682
19	31	70.5	801	5	US-09-815-242-5330
20	31	70.5	1055	5	US-09-815-242-12189
21	31	70.5	4099	5	US-09-897-516-4611
22	30	68.2	94	5	US-09-897-516-6876
23	30	68.2	166	5	US-09-815-242-11445
24	30	68.2	377	5	US-09-708-427-16149
25	30	68.2	381	5	US-09-708-427-16148
26	30	68.2	467	5	US-09-708-427-16147

ALIGNMENTS

RESULT 1

US-09-897-516-6362  
; Sequence 6362, Application US/09897516  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Huesling, Joseph E.  
; APPLICANT: Krasomil-Osterfeld, Karina C.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)B  
; CURRENT APPLICATION NUMBER: US/09/897,516  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/215, 161  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8409  
; SEQ ID NO 6362  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Xenorhabdus sp.  
US-09-897-516-6362

Query Match 77.3%; Score 34; DB 5; Length 219;  
Best Local Similarity 75.0%; Pred.No.16;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
||| |  
Db 42 KPNSKGLY 49

RESULT 2

US-09-815-242-11291  
; Sequence 11291, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes

FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11291  
LENGTH: 380  
TYPE: PRT  
ORGANISM: Helicobacter pylori  
US-09-815-242-11291

Query Match 77.3%; Score 34; DB 5; Length 380;  
Best Local Similarity 75.0%; Pred. No. 28;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNKSLY 8  
| | | | |  
DB 132 KPNTRALY 139

RESULT 3  
US-09-708-427-21872  
Sequence 21872, Application US/09708427  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
TITLE OF INVENTION: THEREBY  
FILE REFERENCE: 2750-1243P  
CURRENT APPLICATION NUMBER: US/09/708,427  
CURRENT FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 21872  
LENGTH: 290  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..290  
OTHER INFORMATION: xaa is any amino acid  
NAME/KEY: misc\_feature  
LOCATION: 1..290  
OTHER INFORMATION: Ceres Seq. ID 1840164  
US-09-708-427-21872

Query Match 72.7%; Score 32; DB 5; Length 290;  
Best Local Similarity 71.4%; Pred. No. 53;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8  
| | | | |  
DB 263 PHDKSIY 269

RESULT 4  
US-09-897-516-5176  
Sequence 5176, Application US/09897516

GENERAL INFORMATION:  
APPLICANT: Corbin, David R.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Huesing, Joseph E.  
APPLICANT: Krasomil-Osterfeld, Karina C.  
APPLICANT: Malvar, Thomas M.  
APPLICANT: Slater, Steven C.  
APPLICANT: Spiridonov, Sergei  
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51847)B  
CURRENT APPLICATION NUMBER: US/09/897,516  
CURRENT FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/215, 161  
PRIOR FILING DATE: 2000-06-30  
NUMBER OF SEQ ID NOS: 8409  
SEQ ID NO 5176  
LENGTH: 300  
TYPE: PRT  
ORGANISM: Xenorhabdus sp.  
US-09-897-516-5176

Query Match 72.7%; Score 32; DB 5; Length 300;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSL 7  
| | | | |  
DB 294 PNDKSL 299

RESULT 5  
US-09-708-427-21871  
Sequence 21871, Application US/09708427  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
TITLE OF INVENTION: THEREBY  
FILE REFERENCE: 2750-1243P  
CURRENT APPLICATION NUMBER: US/09/708,427  
CURRENT FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 21871  
LENGTH: 334  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..334  
OTHER INFORMATION: xaa is any amino acid  
NAME/KEY: misc\_feature  
LOCATION: 1..334  
OTHER INFORMATION: Ceres Seq. ID 1840163  
US-09-708-427-21871

Query Match 72.7%; Score 32; DB 5; Length 334;  
Best Local Similarity 71.4%; Pred. No. 62;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8  
| | | | |  
DB 307 PHDKSIY 313

RESULT 6  
US-09-708-427-21870  
Sequence 21870, Application US/09708427  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID



; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85384  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21870  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..341  
; OTHER INFORMATION: xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..341  
; OTHER INFORMATION: Ceres Seq. ID 1840162  
US-09-708-427-21870

Query Match 72.7%; Score 32; DB 5; Length 341;  
Best Local Similarity 71.4%; Pred. No. 63;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8  
I:||||:  
Db 314 PHDKSIY 320

RESULT 7  
US-09-981-353-122  
; Sequence 122, Application US/09981353  
; GENERAL INFORMATION:  
; APPLICANT: Lasek, Amy W.  
; APPLICANT: Jones, David A.  
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER  
; FILE REFERENCE: PA-0038 US  
; CURRENT APPLICATION NUMBER: US/09/981,353  
; CURRENT FILING DATE: 2001-10-11  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 122  
; LENGTH: 969  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 1344279CD1  
US-09-981-353-122

Query Match 72.7%; Score 32; DB 5; Length 969;  
Best Local Similarity 71.4%; Pred. No. 2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNDKSLY 8  
I:||||:  
Db 95 PNDKGLY 101

RESULT 8  
US-60-337-444-4  
; Sequence 4, Application US/60337444  
; GENERAL INFORMATION:  
; APPLICANT: Tarczynski, Mitchell C.  
; APPLICANT: Olsen, Odd-Arne  
; APPLICANT: Shen, Bo  
; APPLICANT: Lid, Stein E.  
; APPLICANT: Li, Chiangjiang  
; APPLICANT: Jung, Rudolf  
; APPLICANT: Gruis, Darren B.  
; APPLICANT: Ananiev, Evgueni  
; APPLICANT: Nichols, Scott E.

; TITLE OF INVENTION: Methods for Improving Seed and Grain  
; TITLE OF INVENTION: Characteristics  
; FILE REFERENCE: 1390p2  
; CURRENT APPLICATION NUMBER: US/60/337,444  
; CURRENT FILING DATE: 2001-10-25  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 2115  
; TYPE: PRT  
; ORGANISM: arabidopsis  
US-60-337-444-4

Query Match 72.7%; Score 32; DB 7; Length 2115;  
Best Local Similarity 71.4%; Pred. No. 4.6e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8  
I:||||:  
Db 1703 PNDKSLF 1709

RESULT 9  
US-60-337-444-27  
; Sequence 27, Application US/60337444  
; GENERAL INFORMATION:  
; APPLICANT: Tarczynski, Mitchell C.  
; APPLICANT: Olsen, Odd-Arne  
; APPLICANT: Shen, Bo  
; APPLICANT: Lid, Stein E.  
; APPLICANT: Li, Chiangjiang  
; APPLICANT: Jung, Rudolf  
; APPLICANT: Gruis, Darren B.  
; APPLICANT: Ananiev, Evgueni  
; APPLICANT: Nichols, Scott E.  
; TITLE OF INVENTION: Methods for Improving Seed and Grain  
; TITLE OF INVENTION: Characteristics  
; FILE REFERENCE: 1390p2  
; CURRENT APPLICATION NUMBER: US/60/337,444  
; CURRENT FILING DATE: 2001-10-25  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 27  
; LENGTH: 2150  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-60-337-444-27

Query Match 72.7%; Score 32; DB 7; Length 2150;  
Best Local Similarity 71.4%; Pred. No. 4.7e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNDKSLY 8  
I:||||:  
Db 1703 PNDKSLF 1709

RESULT 10  
US-09-637-780B-196  
; Sequence 196, Application US/09637780B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 2750-1096P  
; CURRENT APPLICATION NUMBER: US/09/637,780B  
; CURRENT FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 1401  
; SEQ ID NO 196  
; LENGTH: 220

; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..220  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..220  
; OTHER INFORMATION: Ceres Seq. ID 1481830  
US-09-637-780B-196

Query Match 70.5%; Score 31; DB 5; Length 220;  
Best Local Similarity 62.5%; Pred. No. 62;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
|||  
Db 24 KPADKAIY 31

RESULT 11  
US-09-708-427-72696  
; Sequence 72696, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 72696  
; LENGTH: 220  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..220  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..220  
; OTHER INFORMATION: Ceres Seq. ID 1942391  
US-09-708-427-72696

Query Match 70.5%; Score 31; DB 5; Length 220;  
Best Local Similarity 62.5%; Pred. No. 62;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
|||  
Db 24 KPADKAIY 31

RESULT 12  
US-09-637-780B-195  
; Sequence 195, Application US/09637780B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nickolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1096P  
; CURRENT APPLICATION NUMBER: US/09/637,780B  
; CURRENT FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 1401  
; SEQ ID NO 195  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:

; NAME/KEY: misc\_feature  
; LOCATION: 1..236  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..236  
; OTHER INFORMATION: Ceres Seq. ID 1481829  
US-09-637-780B-195

Query Match 70.5%; Score 31; DB 5; Length 236;  
Best Local Similarity 62.5%; Pred. No. 67;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
|||  
Db 40 KPADKAIY 47

RESULT 13  
US-09-708-427-72695  
; Sequence 72695, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 72695  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..236  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..236  
; OTHER INFORMATION: Ceres Seq. ID 1942390  
US-09-708-427-72695

Query Match 70.5%; Score 31; DB 5; Length 236;  
Best Local Similarity 62.5%; Pred. No. 67;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNDKSLY 8  
|||  
Db 40 KPADKAIY 47

RESULT 14  
US-09-637-780B-194  
; Sequence 194, Application US/09637780B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nickolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1096P  
; CURRENT APPLICATION NUMBER: US/09/637,780B  
; CURRENT FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 1401  
; SEQ ID NO 194  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..291  
; OTHER INFORMATION: Xaa is any amino acid

; NAME/KEY: misc.feature  
; LOCATION: 1..291  
; OTHER INFORMATION: Ceres Seq. ID 1481828  
US-09-637-780B-194

Query Match 70.5%; Score 31; DB 5; Length 291;  
Best Local Similarity 62.5%; Pred. No. 84;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNKSLY 8  
Db 95 KPADKAIY 102

RESULT 15

US-09-708-427-72694  
; Sequence 72694, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 72694  
; LENGTH: 292  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..292  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..292  
; OTHER INFORMATION: Ceres Seq. ID 1942389  
US-09-708-427-72694

Query Match 70.5%; Score 31; DB 5; Length 292;  
Best Local Similarity 62.5%; Pred. No. 84;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPNKSLY 8  
Db 96 KPADKAIY 103

Search completed: January 29, 2002, 10:58:12  
Job time: 2508 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:02 ; Search time 1760.55 seconds  
(without alignments)  
1.419 Million cell updates/sec

Title: US-09-763-397A-9  
Perfect score: 50  
Sequence: 1 KPVIQYDNF 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	9	5	US-08-197-484-21
2	50	100.0	9	6	US-08-278-634-35
3	50	100.0	9	7	US-08-318-856-1
4	50	100.0	9	7	US-08-344-824-40
5	50	100.0	9	7	US-08-344-824A-40
6	50	100.0	9	11	US-08-714-175-21
7	50	100.0	9	11	US-08-714-175A-21
8	50	100.0	9	12	US-08-820-360-21
9	50	100.0	9	18	US-09-454-204A-6

10	50	100.0	9	21	US-09-763-397A-9	Sequence 9, Appl
11	50	100.0	229	18	US-09-454-204A-41	Sequence 41, Appl
12	50	100.0	350	21	US-09-763-397A-2	Sequence 2, Appl
13	39	78.0	329	15	US-09-248-796-15255	Sequence 15255, A
14	38	76.0	464	15	US-09-134-001C-4701	Sequence 4701, Ap
15	38	76.0	464	18	US-09-450-969-6185	Sequence 6185, Ap
16	37	74.0	363	16	US-09-248-796-17364	Sequence 17364, A
17	36	72.0	476	21	US-09-716-504-1	Sequence 1, Appl
18	35	70.0	294	23	US-09-938-294-81	Sequence 81, Appl
19	35	70.0	294	24	US-09-228-466-81	Sequence 81, Appl
20	35	70.0	743	16	US-09-248-796-17817	Sequence 17817, A
21	34	68.0	108	16	US-09-270-767-42539	Sequence 42539, A
22	34	68.0	159	16	US-09-248-796-23065	Sequence 23065, A
23	34	68.0	192	16	US-09-248-796-15447	Sequence 15447, A
24	34	68.0	208	24	US-09-173-464-22140	Sequence 22140, A
25	34	68.0	208	24	US-09-191-637-34180	Sequence 34180, A
26	34	68.0	208	24	US-09-191-681-26881	Sequence 26881, A
27	34	68.0	423	24	US-09-259-128-4965	Sequence 4965, Ap
28	34	68.0	423	24	US-09-314-050-7180	Sequence 7180, Ap
29	34	68.0	542	7	US-08-311-731A-213	Sequence 213, App
30	34	68.0	820	16	US-09-248-796-15055	Sequence 15055, A
31	34	68.0	1223	1	PCT-US01-08631-50317	Sequence 50317, A
32	34	68.0	1393	1	PCT-US01-04098A-1653	Sequence 1653, Ap
33	34	68.0	1398	24	US-09-173-464-26696	Sequence 26696, A
34	34	68.0	1399	24	US-09-191-637-34525	Sequence 34525, A
35	34	68.0	1399	24	US-09-191-681-27139	Sequence 27139, A
36	33	66.0	43	24	US-09-128-476-2791	Sequence 2791, Ap
37	33	66.0	169	1	PCT-US00-03453-2	Sequence 2, Appl
38	33	66.0	169	24	US-09-119-588-9	Sequence 9, Appl
39	33	66.0	180	15	US-09-173-300-28	Sequence 28, Appl
40	33	66.0	185	19	US-09-573-655A-1091	Sequence 1091, Ap
41	33	66.0	264	24	US-09-324-109-29958	Sequence 29958, A
42	33	66.0	265	18	US-09-417-507-28605	Sequence 28605, A
43	33	66.0	317	1	PCT-US01-08631-33758	Sequence 33758, A
44	33	66.0	317	1	PCT-US01-08631-39361	Sequence 39361, A
45	33	66.0	350	1	PCT-US00-03453-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1  
US-08-197-484-21  
; Sequence 21, Application US/08197484  
; GENERAL INFORMATION:  
; APPLICANT: VITIELLO, Maria A.  
; APPLICANT: CHESTNUT, Robert W.  
; APPLICANT: SETTE, Alessandro D.  
; APPLICANT: CELIS, Esteban  
; APPLICANT: GRAY, Howard  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; TITLE OF INVENTION: CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/197.484  
; FILING DATE: 16-FEB-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/935,811  
; FILING DATE: 26-AUG-1992  
; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-197-484-21

Query Match 100.0%; Score 50; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 1 KPIVOYDNF 9

RESULT 2
US-08-763-397a-9
; Sequence 35, Application US/08278634
; GENERAL INFORMATION:
; APPLICANT: SETTE, Alessandro
; APPLICANT: SIDNEY, John
; TITLE OF INVENTION: HLA BINDING PEPTIDES AND THEIR USES
; NUMBER OF SEQUENCES: 288
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Street Tower, 20th
; STREET: Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,634
; FILING DATE: 21-JUL-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/543-9600
; TELEFAX: 415/543-5043
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

```

US-08-278-634-35

Query Match 100.0%; Score 50; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 1 KPIVOYDNF 9

RESULT 3
US-08-318-856-1
; Sequence 1, Application US/08318856
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Peptides of an Antigen, Capable of
; TITLE OF INVENTION: Recognition by or Induction of Cytotoxic T Lymphocytes, and
; TITLE OF INVENTION: Method of Identification
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,856
; FILING DATE: October 3, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 08 068.8
; FILING DATE: April 3, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 17 704.7
; FILING DATE: August 20, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/00711
; FILING DATE: April 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew Jacob
; REGISTRATION NUMBER: 25,154
; REFERENCE/DOCKET NUMBER: 263-PP1R1577US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)-371-8850
; TELEFAX: (202) 371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-08-318-856-1

Query Match 100.0%; Score 50; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9
DB 1 KPIVOYDNF 9

```

RESULT 4  
US-08-344-824-40  
; Sequence 40, Application US/08344824  
; GENERAL INFORMATION:  
; APPLICANT: SETTE, Alessandro  
; TITLE OF INVENTION: HLA BINDING PEPTIDES AND THEIR USES  
; NUMBER OF SEQUENCES: 399  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: One Market Plaza, Steuart Street Tower, 20th  
; STREET: Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/344,824  
; FILING DATE: 23-NOV-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,634  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 14137-80-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-344-824-40

Query Match 100.0%; Score 50; DB 7; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9  
Db 1 KPIVQYDNF 9

RESULT 5  
US-08-344-824A-40  
; Sequence 40, Application US/08344824A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; TITLE OF INVENTION: HLA Binding Peptides and Their Uses  
; NUMBER OF SEQUENCES: 308  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/344,824A  
; FILING DATE: 23-NOV-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,634  
; FILING DATE: 21-JUN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/159,330  
; FILING DATE: 29-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/103,396  
; FILING DATE: 06-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/027,746  
; FILING DATE: 05-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/926,666  
; FILING DATE: 07-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Alicea, Hector A.  
; REGISTRATION NUMBER: 40,891  
; REFERENCE/DOCKET NUMBER: 0118623-0080100S  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-344-824A-40

Query Match 100.0%; Score 50; DB 7; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9  
Db 1 KPIVQYDNF 9

RESULT 6  
US-08-714-175-21  
; Sequence 21, Application US/08714175  
; GENERAL INFORMATION:  
; APPLICANT:  
; APPLICANT:  
; APPLICANT:  
; APPLICANT:  
; APPLICANT:  
; TITLE OF INVENTION: MALARIA PEPTIDES  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WENDEROTH, LIND & PONACK  
; STREET: 805 Fifteenth St., Suite 799,  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: MS-DOS 7  
; SOFTWARE: WordPad for Windows 95  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/714,175

;; FILING DATE: 424  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA: WO PCT/GB95/26982  
;; FILING DATE: 30-MAR-1995  
;; PRIOR APPLICATION DATA: GB 9406492.0  
;; FILING DATE: 31-MAR-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Warren M Cheek, Jr.  
;; REGISTRATION NUMBER: 33,367  
;; REFERENCE/DOCKET NUMBER: 263/KPIR2867US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202)-371-8850  
;; TELEFAX: (202)-371-8856  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-714-175-21

Query Match 100.0%; Score 50; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9  
Db 1 KPIVOYDNF 9

RESULT 7  
US-08-714-175A-21  
; Sequence 21, Application US/08714175A  
; GENERAL INFORMATION:  
; APPLICANT: WENDEROTH, LIND & PONACK  
; APPLICANT: WENDEROTH, LIND & PONACK  
; APPLICANT: WENDEROTH, LIND & PONACK  
; APPLICANT: WENDEROTH, LIND & PONACK  
; TITLE OF INVENTION: MALARIA PEPTIDES  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WENDEROTH, LIND & PONACK  
; STREET: 805 Fifteenth St., Suite 799,  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: MS-DOS 7  
; SOFTWARE: WordPad for Windows 95  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/714,175A  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; PRIOR APPLICATION NUMBER: WO PCT/GB95/26982  
; FILING DATE: 30-MAR-1995  
; PRIOR APPLICATION DATA: GB 9406492.0  
; FILING DATE: 31-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER: 263/KPIR2867US

;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202)-371-8850  
;; TELEFAX: (202)-371-8856  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-714-175A-21

Query Match 100.0%; Score 50; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9  
Db 1 KPIVOYDNF 9

RESULT 8  
US-08-820-360-21  
; Sequence 21, Application US/08820360  
; GENERAL INFORMATION:  
; APPLICANT: VITIELLO, Maria A.  
; APPLICANT: CHESTNUT, Robert W.  
; APPLICANT: SETTE, Alessandro D.  
; APPLICANT: CELIS, Esteban  
; APPLICANT: GRAY, Howard  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; TITLE OF INVENTION: CTL IMMUNITY  
; NUMBER OF SEQUENCES: 154  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/820,360  
; FILING DATE: 12-MAR-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/013,366  
; FILING DATE: 13-MAR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/197,484  
; FILING DATE: 16-FEB-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/935,811  
; FILING DATE: 26-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/874,491  
; FILING DATE: 27-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/827,682  
; FILING DATE: 29-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,568  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; REFERENCE/DOCKET NUMBER: 14137-26-4



; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206) 467-9600  
 ; TELEFAX: (206) 623-6793  
 ; INFORMATION FOR SEQ ID NO: 21:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: peptide  
 US-08-820-360-21

Query Match 100.0%; Score 50; DB 12; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9  
 DB 1 KPIVQYDNF 9

## RESULT 9

US-09-454-204A-6  
 ; Sequence 6, Application US/09454204A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McMichael, Andrew  
 ; APPLICANT: Hill, Adrian V.S.  
 ; APPLICANT: Gilbert, Sarah C.  
 ; APPLICANT: Schneider, Jorg  
 ; APPLICANT: Plebanski, Magdalena  
 ; APPLICANT: Hanke, Tomas  
 ; APPLICANT: Smith, Geoffrey L.  
 ; APPLICANT: Blanchard, Tom  
 ; TITLE OF INVENTION: Methods and Reagents for Vaccination  
 ; FILE OF INVENTION: Which Generate A CD8 T Cell Immune Response  
 ; FILE REFERENCE: 2907.1000-000  
 ; CURRENT APPLICATION NUMBER: US/09/454,204A  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: PCT/GB98/01681  
 ; PRIOR FILING DATE: 1998-06-09  
 ; PRIOR APPLICATION NUMBER: GB 97 11957.2  
 ; NUMBER OF SEQ ID NOS: 78  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Unknown  
 ; FEATURE:  
 ; OTHER INFORMATION: CTL Epitope of the Malaria String  
 US-09-454-204A-6

Query Match 100.0%; Score 50; DB 18; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9  
 DB 1 KPIVQYDNF 9

## RESULT 10

US-09-763-397A-9  
 ; Sequence 9, Application US/09763397A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Government of the United States of America, as represented by the  
 ; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis  
 ; APPLICANT: Control and Prevention  
 ; APPLICANT: Lal, Altaf A.  
 ; APPLICANT: Ping Shi, Ya  
 ; APPLICANT: Hasnain, Seyed E.

; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
 ; FILE REFERENCE: 6395-57049  
 ; CURRENT APPLICATION NUMBER: US/09/763,397A  
 ; CURRENT FILING DATE: 2001-02-16  
 ; PRIOR APPLICATION NUMBER: US 60/097,703  
 ; PRIOR FILING DATE: 1998-08-21  
 ; PRIOR APPLICATION NUMBER: PCT / US99/18869  
 ; PRIOR FILING DATE: 1999-08-19  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 9  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Plasmodium falciparum  
 US-09-763-397A-9

Query Match 100.0%; Score 50; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9  
 DB 1 KPIVQYDNF 9

## RESULT 11

US-09-454-204A-41  
 ; Sequence 41, Application US/09454204A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McMichael, Andrew  
 ; APPLICANT: Hill, Adrian V.S.  
 ; APPLICANT: Gilbert, Sarah C.  
 ; APPLICANT: Schneider, Jorg  
 ; APPLICANT: Plebanski, Magdalena  
 ; APPLICANT: Hanke, Tomas  
 ; APPLICANT: Smith, Geoffrey L.  
 ; APPLICANT: Blanchard, Tom  
 ; TITLE OF INVENTION: Methods and Reagents for Vaccination  
 ; FILE OF INVENTION: Which Generate A CD8 T Cell Immune Response  
 ; FILE REFERENCE: 2907.1000-000  
 ; CURRENT APPLICATION NUMBER: US/09/454,204A  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: PCT/GB98/01681  
 ; PRIOR FILING DATE: 1998-06-09  
 ; PRIOR APPLICATION NUMBER: GB 97 11957.2  
 ; PRIOR FILING DATE: 1997-06-09  
 ; NUMBER OF SEQ ID NOS: 78  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 41  
 ; LENGTH: 229  
 ; TYPE: PRT  
 ; ORGANISM: Unknown  
 ; FEATURE:  
 ; OTHER INFORMATION: Complete Epitope of Malaria String  
 US-09-454-204A-41

Query Match 100.0%; Score 50; DB 18; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9  
 DB 45 KPIVQYDNF 53

## RESULT 12

US-09-763-397A-2  
 ; Sequence 2, Application US/09763397A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Government of the United States of America, as represented by the  
 ; APPLICANT: Secretary of the Department of Health and Human Services, Centers for

; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falciparum  
; FILE REFERENCE: 6395-57049  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Recombinant DNA/Protein  
US-09-763-397A-2

Query Match 100.0%; Score 50; DB 21; Length 350;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVQYDNF 9  
|: |||||  
Db 119 KPIVQYDNF 127

RESULT 13  
US-09-248-796-15255  
; Sequence 15255, Application US/09248796  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796  
; CURRENT FILING DATE: 1999-02-12  
; NUMBER OF SEQ ID NOS: 28206  
; SEQ ID NO 15255  
; LENGTH: 329  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796-15255

Query Match 78.0%; Score 39; DB 16; Length 329;  
Best Local Similarity 75.0%; Pred. No. 35;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVQYDNF 9  
|: |||||  
Db 118 PLAQYDNF 125

RESULT 14  
US-09-134-001C-4701  
; Sequence 4701, Application US/09134001C  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4701  
; LENGTH: 464  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4701

Query Match 76.0%; Score 38; DB 15; Length 464;  
Best Local Similarity 75.0%; Pred. No. 81;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVQYDNF 9  
|: |||||  
Db 348 PIITYDNF 355

RESULT 15  
US-09-450-969-6185  
; Sequence 6185, Application US/09450969  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
; FILE REFERENCE: PATH99-09A  
; CURRENT APPLICATION NUMBER: US/09/450,969  
; CURRENT FILING DATE: 1999-11-29  
; NUMBER OF SEQ ID NOS: 7544  
; SEQ ID NO 6185  
; LENGTH: 464  
; TYPE: PRT  
; ORGANISM: S. epidermidis  
US-09-450-969-6185

Query Match 76.0%; Score 38; DB 18; Length 464;  
Best Local Similarity 75.0%; Pred. No. 81;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVQYDNF 9  
|: |||||  
Db 348 PIITYDNF 355

Search completed: January 29, 2002, 10:56:03  
Job time: 2399 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:11 ; Search time 120.95 Seconds  
(without alignments)  
2.723 Million cell updates/sec

Title: US-09-763-397A-9  
Perfect score: 50  
Sequence: 1 KPVIQYDNF 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	76.0	215	5	US-09-708-427-51402
2	38	76.0	239	5	US-09-708-427-51401
3	38	76.0	267	5	US-09-708-427-51400
4	34	68.0	115	5	US-09-708-427-49931
5	34	68.0	139	5	US-09-708-427-49931
6	33	66.0	355	5	US-09-708-427-54153
7	33	66.0	1328	7	US-60-333-726-1295
8	33	66.0	1328	7	US-60-333-726-1306
9	33	66.0	1328	7	US-60-333-726-1323
10	33	66.0	1328	7	US-60-333-726-1345
11	32	64.0	411	5	US-09-509-234A-47
12	32	64.0	411	5	US-09-509-234C-47
13	32	64.0	420	5	US-09-509-234A-41
14	32	64.0	420	5	US-09-509-234C-41
15	32	64.0	844	5	US-09-815-242-13515
16	32	64.0	1157	5	US-09-708-427-3137
17	32	64.0	1170	5	US-09-708-427-3136
18	32	64.0	1194	5	US-09-708-427-3135
19	31	62.0	152	5	US-09-708-427-48137
20	31	62.0	302	5	US-09-708-427-32838
21	31	62.0	326	1	PCT-US01-27760-641
22	31	62.0	326	1	PCT-US01-27760A-641
23	31	62.0	344	5	US-09-708-427-65943
24	31	62.0	382	5	US-09-708-427-65942
25	31	62.0	409	5	US-09-897-516-5513
26	31	62.0	679	5	US-09-976-800-83

ALIGNMENTS

RESULT 1

US-09-708-427-51402  
; Sequence 51402, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 51402  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..215  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..215  
; OTHER INFORMATION: Ceres Seq. ID 1926741  
; US-09-708-427-51402

Query Match 76.0%; Score 38; DB 5; Length 215;  
Best Local Similarity 75.0%; Pred.No. 2.5;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPVIQYDYN 8  
||:|:|:|:  
Db 82 KPLVKYDYN 89

RESULT 2

US-09-708-427-51401  
; Sequence 51401, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 51401  
; LENGTH: 239

Sequence 84, Appl  
Sequence 117, App  
Sequence 118, App  
Sequence 14, Appl  
Sequence 16, Appl  
Sequence 31790, A  
Sequence 1339, Ap  
Sequence 15, Appl  
Sequence 31789, A  
Sequence 31788, A  
Sequence 36, Appl  
Sequence 2171, Ap  
Sequence 7302, Ap  
Sequence 2170, Ap  
Sequence 625, App  
Sequence 625, App  
Sequence 51, Appl  
Sequence 51, Appl  
Sequence 7448, Ap

679 5 US-09-976-800-84  
679 5 US-09-976-800-117  
679 5 US-09-976-800-118  
993 5 US-09-635-949-14  
994 5 US-09-635-949-16  
183 5 US-09-708-427-31790  
185 7 US-60-333-726-1339  
195 5 US-09-868-352-15  
223 5 US-09-708-427-31789  
270 5 US-09-708-427-31788  
310 5 US-09-265-585C-36  
335 5 US-09-620-394B-2171  
339 5 US-09-897-516-7302  
359 5 US-09-620-394B-2170  
366 1 PCT-US01-27760-625  
366 1 PCT-US01-27760A-625  
410 5 US-09-509-234A-51  
410 5 US-09-509-234C-51  
427 5 US-09-708-427-7448

31 62.0  
31 62.0  
31 62.0  
31 62.0  
31 62.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0  
30 60.0

27  
28  
29  
30  
31  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45

; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..239  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..239  
; OTHER INFORMATION: Ceres Seq. ID 1926740  
; US-09-708-427-51401

Query Match 76.0%; Score 38; DB 5; Length 239;  
Best Local Similarity 75.0%; Pred. No. 2.8;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDN 8  
||:|:|  
Db 106 KPLVKYDN 113

RESULT 3  
US-09-708-427-51400  
; Sequence 51400, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 51400  
; LENGTH: 267  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..267  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..267  
; OTHER INFORMATION: Ceres Seq. ID 1926739  
; US-09-708-427-51400

Query Match 76.0%; Score 38; DB 5; Length 267;  
Best Local Similarity 75.0%; Pred. No. 3.2;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYDN 8  
||:|:|  
Db 134 KPLVKYDN 141

RESULT 4  
US-09-708-427-49931  
; Sequence 49931, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 49931  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:

; NAME/KEY: misc\_feature  
; LOCATION: 1..115  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..115  
; OTHER INFORMATION: Ceres Seq. ID 1923418  
; US-09-708-427-49931

Query Match 68.0%; Score 34; DB 5; Length 115;  
Best Local Similarity 55.6%; Pred. No. 7.7;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9  
||:|:|  
Db 94 KPLVGYDDF 102

RESULT 5  
US-09-708-427-49343  
; Sequence 49343, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 49343  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..139  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..139  
; OTHER INFORMATION: Ceres Seq. ID 1921669  
; US-09-708-427-49343

Query Match 68.0%; Score 34; DB 5; Length 139;  
Best Local Similarity 55.6%; Pred. No. 9.6;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPIVOYDNF 9  
||:|:|  
Db 48 KPLVGYDDF 56

RESULT 6  
US-09-708-427-54153  
; Sequence 54153, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 54153  
; LENGTH: 355  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..355  
; OTHER INFORMATION: Xaa is any amino acid

; NAME/KEY: misc\_feature  
; LOCATION: 1..355  
; OTHER INFORMATION: Ceres Seq. ID 1932726  
US-09-708-427-54153

Query Match 66.0%; Score 33; DB 5; Length 355;  
Best Local Similarity 62.5%; Pred. No. 45;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPIVQYDN 8  
Db 101 KPLLYDN 108

RESULT 7  
US-60-333-726-1295  
; Sequence 1295, Application US/60333726  
; GENERAL INFORMATION:  
; APPLICANT: Qiangdong Zeng et al.  
; TITLE OF INVENTION: Systemic Discovery of New Genes  
; FILE REFERENCE: 032796-090  
; CURRENT APPLICATION NUMBER: US/60/333,726  
; CURRENT FILING DATE: 2001-11-29  
; NUMBER OF SEQ ID NOS: 1346  
; SEQ ID NO 1295  
; LENGTH: 1328  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-60-333-726-1295

Query Match 66.0%; Score 33; DB 7; Length 1328;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVOYDNF 9  
Db 129 PIVOYDNF 136

RESULT 8  
US-60-333-726-1306  
; Sequence 1306, Application US/60333726  
; GENERAL INFORMATION:  
; APPLICANT: Qiangdong Zeng et al.  
; TITLE OF INVENTION: Systemic Discovery of New Genes  
; FILE REFERENCE: 032796-090  
; CURRENT APPLICATION NUMBER: US/60/333,726  
; CURRENT FILING DATE: 2001-11-29  
; NUMBER OF SEQ ID NOS: 1346  
; SEQ ID NO 1306  
; LENGTH: 1328  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-60-333-726-1306

Query Match 66.0%; Score 33; DB 7; Length 1328;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVOYDNF 9  
Db 129 PIVOYDNF 136

RESULT 9  
US-60-333-726-1323  
; Sequence 1323, Application US/60333726  
; GENERAL INFORMATION:  
; APPLICANT: Qiangdong Zeng et al.  
; TITLE OF INVENTION: Systemic Discovery of New Genes

; FILE REFERENCE: 032796-090  
; CURRENT APPLICATION NUMBER: US/60/333,726  
; CURRENT FILING DATE: 2001-11-29  
; NUMBER OF SEQ ID NOS: 1346  
; SEQ ID NO 1323  
; LENGTH: 1328  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-60-333-726-1323

Query Match 66.0%; Score 33; DB 7; Length 1328;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVOYDNF 9  
Db 129 PIVOYDNF 136

RESULT 10  
US-60-333-726-1345  
; Sequence 1345, Application US/60333726  
; GENERAL INFORMATION:  
; APPLICANT: Qiangdong Zeng et al.  
; TITLE OF INVENTION: Systemic Discovery of New Genes  
; FILE REFERENCE: 032796-090  
; CURRENT APPLICATION NUMBER: US/60/333,726  
; CURRENT FILING DATE: 2001-11-29  
; NUMBER OF SEQ ID NOS: 1346  
; SEQ ID NO 1345  
; LENGTH: 1328  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-60-333-726-1345

Query Match 66.0%; Score 33; DB 7; Length 1328;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVOYDNF 9  
Db 129 PIVOYDNF 136

RESULT 11  
US-09-509-234A-47  
; Sequence 47, Application US/09509234A  
; GENERAL INFORMATION:  
; APPLICANT: Vannuffel, Pascal  
; APPLICANT: Gala, Jean-Luc  
; TITLE OF INVENTION: GENETIC SEOS.....METHODS & DEVICES FOR  
; FILE REFERENCE: DECLE27.001APC  
; CURRENT APPLICATION NUMBER: US/09/509,234A  
; CURRENT FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: EP 97870146.4  
; PRIOR FILING DATE: 1997-09-26  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 47  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Staphylococcus capitis fema  
US-09-509-234A-47

Query Match 64.0%; Score 32; DB 5; Length 411;  
Best Local Similarity 57.1%; Pred. No. 85;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVOYDN 8

Db 70 PVIDYDN 76

RESULT 12  
US-09-509-234C-47  
; Sequence 47, Application US/09509234C  
; GENERAL INFORMATION:  
; APPLICANT: Vannuffel, Pascal  
; APPLICANT: Gala, Jean-Luc  
; TITLE OF INVENTION: GENETIC SEQS.....METHODS & DEVICES FOR  
; TITLE OF INVENTION: ID. OF STAPHYLOCOCCI STRAINS  
; FILE REFERENCE: DECLE27.001APC  
; CURRENT APPLICATION NUMBER: US/09/509,234C  
; CURRENT FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: EP 97870146.4  
; PRIOR FILING DATE: 1997-09-26  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 47  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Staphylococcus capitis femA  
US-09-509-234C-47

Query Match 64.0%; Score 32; DB 5; Length 411;  
Best Local Similarity 57.1%; Pred. No. 85;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVQYDN 8  
I:: III  
Db 70 PVIDYDN 76

RESULT 13  
US-09-509-234A-41  
; Sequence 41, Application US/09509234A  
; GENERAL INFORMATION:  
; APPLICANT: Vannuffel, Pascal  
; APPLICANT: Gala, Jean-Luc  
; TITLE OF INVENTION: GENETIC SEQS.....METHODS & DEVICES FOR  
; TITLE OF INVENTION: ID. OF STAPHYLOCOCCI STRAINS  
; FILE REFERENCE: DECLE27.001APC  
; CURRENT APPLICATION NUMBER: US/09/509,234A  
; CURRENT FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: EP 97870146.4  
; PRIOR FILING DATE: 1997-09-26  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 420  
; TYPE: PRT  
; ORGANISM: Staphylococcus haemolyticus femA  
US-09-509-234A-41

Query Match 64.0%; Score 32; DB 5; Length 420;  
Best Local Similarity 57.1%; Pred. No. 87;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVQYDN 8  
I:: III  
Db 76 PVIDYDN 82

RESULT 14  
US-09-509-234C-41  
; Sequence 41, Application US/09509234C  
; GENERAL INFORMATION:  
; APPLICANT: Vannuffel, Pascal  
; APPLICANT: Gala, Jean-Luc  
; TITLE OF INVENTION: GENETIC SEQS.....METHODS & DEVICES FOR

; TITLE OF INVENTION: ID. OF STAPHYLOCOCCI STRAINS  
; FILE REFERENCE: DECLE27.001APC  
; CURRENT APPLICATION NUMBER: US/09/509,234C  
; CURRENT FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: EP 97870146.4  
; PRIOR FILING DATE: 1997-09-26  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 420  
; TYPE: PRT  
; ORGANISM: Staphylococcus haemolyticus femA  
US-09-509-234C-41

Query Match 64.0%; Score 32; DB 5; Length 420;  
Best Local Similarity 57.1%; Pred. No. 87;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVQYDN 8  
I:: III  
Db 76 PVIDYDN 82

RESULT 15  
US-09-815-242-13515  
; Sequence 13515, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 13515  
; LENGTH: 844  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13515

Query Match 64.0%; Score 32; DB 5; Length 844;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPIVOYD 7  
I::: III  
Db 239 KPIVYD 245



Search completed: January 29, 2002, 10:58:11  
Job time: 2507 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:02 : Search time 1760.55 Seconds  
(without alignments)  
1.419 Million cell updates/sec

Title: US-09-763-397A-8

Perfect score: 59

Sequence: 1 WSPCSVTGC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/PCRTUS\_COMB.pep:\*

2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*

3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*

4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep:\*

5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*

6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*

7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*

8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*

9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*

10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*

11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*

12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*

13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*

14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*

15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*

16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*

17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*

18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*

19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*

20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*

21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*

22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*

23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*

24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by the chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	59	100.0	9	15	US-09-197-770-2
2	59	100.0	9	21	US-09-763-397A-8
3	59	100.0	18	3	US-07-947-033-1
4	59	100.0	18	3	US-07-947-033-7
5	59	100.0	18	3	US-07-947-033-8
6	59	100.0	18	3	US-07-947-033-9
7	59	100.0	18	3	US-07-947-033-10
8	59	100.0	18	5	US-08-119-694-1
9	59	100.0	18	5	US-08-119-694-7

10	59	100.0	18	5	US-08-119-694-8	Sequence 8, Appl1
11	59	100.0	18	5	US-08-119-694-9	Sequence 9, Appl1
12	59	100.0	18	5	US-08-119-694-10	Sequence 10, Appl1
13	59	100.0	18	5	US-08-119-694B-1	Sequence 1, Appl1
14	59	100.0	18	5	US-08-119-694B-7	Sequence 7, Appl1
15	59	100.0	18	5	US-08-119-694B-8	Sequence 8, Appl1
16	59	100.0	18	5	US-08-119-694B-9	Sequence 9, Appl1
17	59	100.0	18	5	US-08-119-694B-10	Sequence 10, Appl1
18	59	100.0	18	7	US-08-395-602-1	Sequence 1, Appl1
19	59	100.0	20	18	US-09-454-204A-28	Sequence 28, Appl1
20	59	100.0	21	24	US-60-143-992-856	Sequence 856, App
21	59	100.0	23	1	PCT-US98-25922-9	Sequence 9, Appl1
22	59	100.0	23	1	PCT-US98-25922-33	Sequence 33, Appl1
23	59	100.0	23	7	US-08-395-602-2	Sequence 2, Appl1
24	59	100.0	27	24	US-60-138-684-1035	Sequence 1035, Ap
25	59	100.0	33	1	PCT-US98-25922-34	Sequence 34, Appl1
26	59	100.0	54	1	PCT-US99-31025-46	Sequence 46, Appl1
27	59	100.0	54	18	US-09-471-179-46	Sequence 46, Appl1
28	59	100.0	54	21	US-09-782-952-5	Sequence 5, Appl1
29	59	100.0	54	24	US-60-258-373-6	Sequence 6, Appl1
30	59	100.0	54	24	US-60-297-863-6	Sequence 6, Appl1
31	59	100.0	55	3	US-07-947-033-2	Sequence 2, Appl1
32	59	100.0	55	3	US-07-947-033-3	Sequence 3, Appl1
33	59	100.0	55	3	US-07-947-033-4	Sequence 4, Appl1
34	59	100.0	55	3	US-07-947-033-5	Sequence 5, Appl1
35	59	100.0	55	5	US-08-119-694-2	Sequence 2, Appl1
36	59	100.0	55	5	US-08-119-694-3	Sequence 3, Appl1
37	59	100.0	55	5	US-08-119-694-4	Sequence 4, Appl1
38	59	100.0	55	5	US-08-119-694-5	Sequence 5, Appl1
39	59	100.0	55	5	US-08-119-694B-2	Sequence 2, Appl1
40	59	100.0	55	5	US-08-119-694B-3	Sequence 3, Appl1
41	59	100.0	55	5	US-08-119-694B-4	Sequence 4, Appl1
42	59	100.0	55	5	US-08-119-694B-5	Sequence 5, Appl1
43	59	100.0	172	24	US-60-143-992-1057	Sequence 1057, Ap
44	59	100.0	229	18	US-09-454-204A-41	Sequence 41, Appl1
45	59	100.0	350	21	US-09-763-397A-2	Sequence 2, Appl1

#### ALIGNMENTS

RESULT 1

US-09-197-770-2

Sequence 2, Application US/09197770A

GENERAL INFORMATION:

APPLICANT: Tuszyński, George

APPLICANT: Williams, Talfy

APPLICANT: Actor, Paul

TITLE OF INVENTION: RETROINVERSO POLYPEPTIDES THAT MIMIC OR INHIBIT

FILE REFERENCE: 07206-0021

CURRENT APPLICATION NUMBER: US/09/197,770A

CURRENT FILING DATE: 1998-11-23

NUMBER OF SEQ ID NOS: 38

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 9

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-09-197-770-2

Query Match 100.0%; Score 59; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTGC 9  
DB 1 WSPCSVTGC 9

RESULT 2  
US-09-763-397A-8  
; Sequence 8, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; PRIOR FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-8

Query Match 100.0%; Score 59; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.9e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTGC 9  
DB 1 WSPCSVTGC 9

RESULT 3  
US-07-947-033-1  
; Sequence 1, Application US/07947033  
; GENERAL INFORMATION:  
; APPLICANT: Cerami, Carla  
; APPLICANT: Frevert, Ute  
; APPLICANT: Simnis, Photini  
; APPLICANT: Nussenzweig, Victor  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING  
; TITLE OF INVENTION: HEPATOCYTE INVASION BY MALARIAL SPOROZOITES  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Darby & Darby, P.C.  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage  
; COMPUTER: AST Premium II 386/33  
; OPERATING SYSTEM: DOS 3.3  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/947,033  
; FILING DATE: 19920917  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adda C. Gogoris, Esq.  
; REGISTRATION NUMBER: 29,714  
; REFERENCE/DOCKET NUMBER: 5986/07607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 527-7700  
; TELEFAX: (212) 753-6237  
; TELEX: 236687  
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
US-07-947-033-1

Query Match 100.0%; Score 59; DB 3; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTGC 9  
DB 2 WSPCSVTGC 10

RESULT 4  
US-07-947-033-7  
; Sequence 7, Application US/07947033  
; GENERAL INFORMATION:  
; APPLICANT: Cerami, Carla  
; APPLICANT: Frevert, Ute  
; APPLICANT: Simnis, Photini  
; APPLICANT: Nussenzweig, Victor  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING  
; TITLE OF INVENTION: HEPATOCYTE INVASION BY MALARIAL SPOROZOITES  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Darby & Darby, P.C.  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage  
; COMPUTER: AST Premium II 386/33  
; OPERATING SYSTEM: DOS 3.3  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/947,033  
; FILING DATE: 19920917  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adda C. Gogoris, Esq.  
; REGISTRATION NUMBER: 29,714  
; REFERENCE/DOCKET NUMBER: 5986/07607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 527-7700  
; TELEFAX: (212) 753-6237  
; TELEX: 236687  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
US-07-947-033-7

Query Match 100.0%; Score 59; DB 3; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTGC 9  
DB 2 WSPCSVTGC 10

RESULT 5  
US-07-947-033-8  
; Sequence 8, Application US/07947033  
; GENERAL INFORMATION:  
; APPLICANT: Cerami, Carla

```

: APPLICANT: Frevert, Ute
: APPLICANT: Siminis, Photini
: APPLICANT: Nussenzweig, Victor
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Darby & Darby, P.C.
: STREET: 805 Third Avenue
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10022
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage
: OPERATING SYSTEM: DOS 3.3
: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/947,033
: FILING DATE: 19920917
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Adda C. Gogoris, Esq.
: REGISTRATION NUMBER: 29,714
: REFERENCE/DOCKET NUMBER: 5986/07607
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 527-7700
: TELEFAX: (212) 753-6237
: TELEX: 236687
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 18 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: unknown
: US-07-947-033-8

Query Match      100.0%; Score 59; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTGC 9
   |||||
Db 2 WSPCSVTGC 10

RESULT 6
US-07-947-033-9
: Sequence 9, Application US/07947033
: GENERAL INFORMATION:
: APPLICANT: Cerami, Carla
: APPLICANT: Frevert, Ute
: APPLICANT: Siminis, Photini
: APPLICANT: Nussenzweig, Victor
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Darby & Darby, P.C.
: STREET: 805 Third Avenue
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10022
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage
: OPERATING SYSTEM: DOS 3.3
: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/947,033
: FILING DATE: 19920917

```

```

: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Adda C. Gogoris, Esq.
: REGISTRATION NUMBER: 29,714
: REFERENCE/DOCKET NUMBER: 5986/07607
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 527-7700
: TELEFAX: (212) 753-6237
: TELEX: 236687
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 18 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: unknown
: US-07-947-033-9

Query Match      100.0%; Score 59; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTGC 9
   |||||
Db 2 WSPCSVTGC 10

RESULT 7
US-07-947-033-10
: Sequence 10, Application US/07947033
: GENERAL INFORMATION:
: APPLICANT: Cerami, Carla
: APPLICANT: Frevert, Ute
: APPLICANT: Siminis, Photini
: APPLICANT: Nussenzweig, Victor
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Darby & Darby, P.C.
: STREET: 805 Third Avenue
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10022
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage
: OPERATING SYSTEM: DOS 3.3
: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/947,033
: FILING DATE: 19920917
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Adda C. Gogoris, Esq.
: REGISTRATION NUMBER: 29,714
: REFERENCE/DOCKET NUMBER: 5986/07607
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 527-7700
: TELEFAX: (212) 753-6237
: TELEX: 236687
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 18 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: unknown
: US-07-947-033-10

Query Match      100.0%; Score 59; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 WSPCSVTCG 9  
Db 2 WSPCSVTCG 10

## RESULT 8

US-08-119-694-1  
Sequence 1, Application US/08119694  
GENERAL INFORMATION:  
APPLICANT: Cerami, Carla  
APPLICANT: Frevert, Ute  
APPLICANT: Simins, Photini  
APPLICANT: Nussenzweig, Victor  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING  
TITLE OF INVENTION: HEPATOCYTE INVASION BY MALARIAL  
SPOROZOITES  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Darby & Darby, P.C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage  
OPERATING SYSTEM: DOS 3.3  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,694  
FILING DATE: 10-Sept-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Adda C. Gogoris, Esq.  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17607-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
US-08-119-694-1

Query Match 100.0%; Score 59; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSPCSVTCG 9  
Db 2 WSPCSVTCG 10

## RESULT 9

US-08-119-694-7  
Sequence 7, Application US/08119694  
GENERAL INFORMATION:  
APPLICANT: Cerami, Carla  
APPLICANT: Frevert, Ute  
APPLICANT: Simins, Photini  
APPLICANT: Nussenzweig, Victor  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING  
TITLE OF INVENTION: HEPATOCYTE INVASION BY MALARIAL  
SPOROZOITES  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Darby & Darby, P.C.  
STREET: 805 Third Avenue  
CITY: New York

STATE: New York  
COUNTRY: United States of America  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage  
COMPUTER: AST Premium II 386/33  
OPERATING SYSTEM: DOS 3.3  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,694  
FILING DATE: 10-Sept-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Adda C. Gogoris, Esq.  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17607-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
US-08-119-694-7

Query Match 100.0%; Score 59; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSPCSVTCG 9  
Db 2 WSPCSVTCG 10

RESULT 10  
US-08-119-694-8  
Sequence 8, Application US/08119694  
GENERAL INFORMATION:  
APPLICANT: Cerami, Carla  
APPLICANT: Frevert, Ute  
APPLICANT: Simins, Photini  
APPLICANT: Nussenzweig, Victor  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING  
TITLE OF INVENTION: HEPATOCYTE INVASION BY MALARIAL  
SPOROZOITES  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Darby & Darby, P.C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage  
COMPUTER: AST Premium II 386/33  
OPERATING SYSTEM: DOS 3.3  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,694  
FILING DATE: 10-Sept-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Adda C. Gogoris, Esq.  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17607-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
US-08-119-694-8

Query Match 100.0%; Score 59; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSPCSVTCG 9  
DB 2 WSPCSVTCG 10

RESULT 11  
US-08-119-694-9  
Sequence 9, Application US/08119694  
GENERAL INFORMATION:  
APPLICANT: Cerami, Carla  
APPLICANT: Frevert, Ute  
APPLICANT: Sinnis, Photini  
APPLICANT: Nussenzweig, Victor  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING  
TITLE OF INVENTION: HEPATOCYTE INVASION BY MALARIAL  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby, P. C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage  
COMPUTER: AST Premium II 386/33  
OPERATING SYSTEM: DOS 3.3  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,694  
FILING DATE: 10-Sept-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Adda C. Gogoris, Esq.  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17607-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
US-08-119-694-9

Query Match 100.0%; Score 59; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSPCSVTCG 9  
DB 2 WSPCSVTCG 10

RESULT 12  
US-08-119-694-10  
Sequence 10, Application US/08119694  
GENERAL INFORMATION:  
APPLICANT: Cerami, Carla  
APPLICANT: Frevert, Ute

APPLICANT: Sinnis, Photini  
APPLICANT: Nussenzweig, Victor  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING  
TITLE OF INVENTION: HEPATOCYTE INVASION BY MALARIAL  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby, P. C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage  
COMPUTER: AST Premium II 386/33  
OPERATING SYSTEM: DOS 3.3  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,694  
FILING DATE: 10-Sept-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Adda C. Gogoris, Esq.  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17607-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
US-08-119-694-10

Query Match 100.0%; Score 59; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSPCSVTCG 9  
DB 2 WSPCSVTCG 10

RESULT 13  
US-08-119-694B-1  
Sequence 1, Application US/08119694B  
GENERAL INFORMATION:  
APPLICANT: Cerami, Carla  
APPLICANT: Frevert, Ute  
APPLICANT: Sinnis, Photini  
APPLICANT: Nussenzweig, Victor  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: INHIBITING HEPATOCYTE INVASION  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby, P. C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5, inch, 360 KB storage  
COMPUTER: COMPAQ PROLINEA MT 4/66  
OPERATING SYSTEM: DOS 3.3  
SOFTWARE: Word Perfect 5.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,694B  
FILING DATE: 10-Sept-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Adda C. Gogoris, Esq.  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17607-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
US-08-119-694B-1

Query Match 100.0%; Score 59; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTG 9  
Db 2 WSPCSVTG 10

RESULT 14  
US-08-119-694B-7  
Sequence 7, Application US/08119694B  
GENERAL INFORMATION:  
APPLICANT: Cerami, Carla  
APPLICANT: Frevert, Ute  
APPLICANT: Siminis, Photini  
APPLICANT: Nussenzweig, Victor  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: INHIBITING HEPATOCYTE INVASION  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby, P.C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5, inch, 360 KB storage  
COMPUTER: COMPAQ PROLINEA MT 4/66  
OPERATING SYSTEM: DOS 3.3  
SOFTWARE: Word Perfect 5.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,694B  
FILING DATE: 10-Sept-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Adda C. Gogoris, Esq.  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17607-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
US-08-119-694B-7

Query Match 100.0%; Score 59; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTG 9

Db 2 WSPCSVTG 10

RESULT 15  
US-08-119-694B-8  
Sequence 8, Application US/08119694B  
GENERAL INFORMATION:  
APPLICANT: Cerami, Carla  
APPLICANT: Frevert, Ute  
APPLICANT: Siminis, Photini  
APPLICANT: Nussenzweig, Victor  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: INHIBITING HEPATOCYTE INVASION  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby, P.C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5, inch, 360 KB storage  
COMPUTER: COMPAQ PROLINEA MT 4/66  
OPERATING SYSTEM: DOS 3.3  
SOFTWARE: Word Perfect 5.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,694B  
FILING DATE: 10-Sept-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Adda C. Gogoris, Esq.  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17607-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
US-08-119-694B-8

Query Match 100.0%; Score 59; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSVTG 9  
Db 2 WSPCSVTG 10

Search completed: January 29, 2002, 10:56:02  
Job time: 2398 sec



---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:10 : Search time 120.95 Seconds  
(without alignments)  
2.723 Million cell updates/sec

Title: US-09-763-397a-8

Perfect score: 59

Sequence: 1 WSPCSVTGC 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PC7\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	9	US-09-197-770B-2	Sequence 2, Appli
2	59	100.0	9	US-09-462-909B-2	Sequence 2, Appli
3	59	100.0	18	US-09-980-564-13	Sequence 13, Appli
4	59	100.0	54	US-09-858-068-6	Sequence 6, Appli
5	59	100.0	54	US-09-858-081-6	Sequence 6, Appli
6	59	100.0	54	US-10-014-070-9	Sequence 9, Appli
7	56	94.9	9	US-09-197-770B-17	Sequence 17, Appli
8	56	94.9	525	US-09-978-189-301	Sequence 301, App
9	56	94.9	525	US-09-978-192-301	Sequence 301, App
10	56	94.9	525	US-09-978-687-301	Sequence 301, App
11	56	94.9	525	US-09-978-824-301	Sequence 301, App
12	54	91.5	787	US-09-970-966-207	Sequence 207, App
13	54	91.5	807	US-09-970-966-186	Sequence 186, App
14	54	91.5	807	US-09-611-526-3833	Sequence 3833, Ap
15	53	89.8	9	US-09-197-770B-14	Sequence 14, Appli
16	51	86.4	9	US-09-197-770B-9	Sequence 9, Appli
17	51	86.4	9	US-09-197-770B-15	Sequence 15, Appli
18	51	86.4	9	US-09-462-909B-3	Sequence 4, Appli
19	51	86.4	9	US-09-462-909B-4	Sequence 4, Appli
20	51	86.4	9	US-09-462-909B-11	Sequence 11, Appli
21	51	86.4	23	US-09-197-770B-12	Sequence 12, Appli
22	51	86.4	57	US-09-802-094-5	Sequence 5, Appli
23	51	86.4	57	US-09-802-094-6	Sequence 6, Appli
24	51	86.4	57	US-09-989-687-7	Sequence 7, Appli
25	51	86.4	57	US-09-989-687-10	Sequence 10, Appli
26	51	86.4	242	US-09-919-603-5	Sequence 5, Appli

27	51	86.4	300	5	US-09-919-603-7	Sequence 7, Appli
28	51	86.4	367	6	US-10-001-054-50	Sequence 50, Appli
29	51	86.4	926	5	US-09-715-417A-28	Sequence 28, Appli
30	51	86.4	939	5	US-09-854-845-16	Sequence 16, Appli
31	51	86.4	954	5	US-09-854-845-14	Sequence 14, Appli
32	51	86.4	1034	5	US-09-854-845-6	Sequence 6, Appli
33	51	86.4	1049	5	US-09-854-845-2	Sequence 2, Appli
34	51	86.4	1078	5	US-09-854-845-8	Sequence 8, Appli
35	51	86.4	1093	5	US-09-854-845-4	Sequence 4, Appli
36	51	86.4	1136	5	US-09-854-845-12	Sequence 12, Appli
37	51	86.4	1151	5	US-09-854-845-10	Sequence 10, Appli
38	51	86.4	1152	5	US-09-919-603-1	Sequence 1, Appli
39	51	86.4	1168	5	US-09-919-603-2	Sequence 2, Appli
40	51	86.4	1170	6	US-10-008-093-2	Sequence 2, Appli
41	51	86.4	1170	6	US-10-020-141-12	Sequence 12, Appli
42	51	86.4	1172	6	US-10-008-093-4	Sequence 4, Appli
43	51	86.4	1172	6	US-10-020-141-14	Sequence 14, Appli
44	51	86.4	1203	5	US-09-989-687-5	Sequence 5, Appli
45	50	84.7	465	5	US-09-969-515-12	Sequence 12, Appli

#### ALIGNMENTS

```
RESULT 1
US-09-197-770B-2
: Sequence 2, Application US/09197770B
: GENERAL INFORMATION:
: APPLICANT: Tuszyński, George
: APPLICANT: Williams, Taffy
: APPLICANT: Actor, Paul
: TITLE OF INVENTION: RETROINVERSO POLYPEPTIDES THAT MIMIC OR INHIBIT
: FILE REFERENCE: 07206-0021
: CURRENT FILING DATE: US/09/197,770B
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 9
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: synthetic
: OTHER INFORMATION: fragment/ analog of thrombospondin
US-09-197-770B-2
:
: Query Match 100.0%; Score 59; DB 5; Length 9;
: Best Local Similarity 100.0%; Pred. No. 1.6e+05;
: Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WSPCSVTGC 9
DB 1 WSPCSVTGC 9
RESULT 2
US-09-462-909B-2
: Sequence 2, Application US/09462909B
: GENERAL INFORMATION:
: APPLICANT: MEINEL, Annie
: APPLICANT: MONNERIE, Hubert
: APPLICANT: GOBRON, Stephane
: TITLE OF INVENTION: NOVEL PEPTIDES AND POLYPEPTIDES USEFUL FOR REGENERATING THE NE
: FILE REFERENCE: 065691/0179
: CURRENT FILING DATE: 2000-02-14
: PRIOR APPLICATION NUMBER: US/09/462,909B
: PRIOR FILING DATE: 1998-07-16
: PRIOR APPLICATION NUMBER: FR 97/09016
: PRIOR FILING DATE: 1997-07-16
```

```
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Bovine
US-09-462-909b-2
```

```
Query Match          100.0%; Score 59; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WSPCSVTGC 9
    |||||
Db 1 WSPCSVTGC 9
```

```
RESULT 3
US-09-980-564-13
```

```
; Sequence 13, Application US/09980564
; GENERAL INFORMATION:
; APPLICANT: University of Washington
; TITLE OF INVENTION: RECOMBINANT ADENOVIIRAL VECTORS FOR CELL SPECIFIC
; TITLE OF INVENTION: INFECTION AND GENOME INTEGRATION AND EXPRESSING
; TITLE OF INVENTION: CHIMERIC FIBER PROTEINS
; FILE REFERENCE: 30429.2w001
; CURRENT APPLICATION NUMBER: US/09/980,564
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/137,213
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/161,097
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; PUBLICATION INFORMATION:
; AUTHORS: Cerami, C.
; JOURNAL: Cell
; VOLUME: 70
; PAGES: 1021-33
; DATE: 1992
; PUBLICATION INFORMATION:
; AUTHORS: Chatterjee, S.
; JOURNAL: Infect. Immun.
; VOLUME: 63
; PAGES: 4375-81
; DATE: 1995
US-09-980-564-13
```

```
Query Match          100.0%; Score 59; DB 5; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WSPCSVTGC 9
    |||||
Db 2 WSPCSVTGC 10
```

```
RESULT 4
US-09-858-068-6
```

```
; Sequence 6, Application US/09858068
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Cook, William James
; APPLICANT: Sllós-Santlago, Imaculada
; TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-057001
```

```
; CURRENT APPLICATION NUMBER: US/09/858,068
; CURRENT FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/204,159
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: 60/204,160
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-09-858-068-6
```

```
Query Match          100.0%; Score 59; DB 5; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WSPCSVTGC 9
    |||||
Db 6 WSPCSVTGC 14
```

```
RESULT 5
US-09-858-081-6
```

```
; Sequence 6, Application US/09858081
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Cook, William James
; APPLICANT: Sllós-Santlago, Imaculada
; TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-049001
; CURRENT APPLICATION NUMBER: US/09/858,081
; CURRENT FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/204,159
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: 60/204,160
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-09-858-081-6
```

```
Query Match          100.0%; Score 59; DB 5; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WSPCSVTGC 9
    |||||
Db 6 WSPCSVTGC 14
```

```
RESULT 6
US-10-014-070-9
```

```
; Sequence 9, Application US/10014070
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhara
; APPLICANT: Curtiss, Rory A.J.
; APPLICANT: Spurling, Heidi Lynn
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 53014, A Human Metalloprotease Family
; TITLE OF INVENTION: Member and Uses Therefor
; FILE REFERENCE: MPI2000-523PIRCP1(M)
```

```

: CURRENT APPLICATION NUMBER: US/10/014,070
: CURRENT FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: 60/258,373
: PRIOR FILING DATE: 2000-12-22
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 54
: TYPE: PRP
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: consensus
: US-10-014-070-9

Query Match
Best Local Similarity 100.0%; Score 59; DB 6; Length 54;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WSPCSVTGC 9
   |||||
Db 6 WSPCSVTGC 14

-RESULT 7
: US-09-197-770B-17
: Sequence 17, Application US/09197770B
: GENERAL INFORMATION:
: APPLICANT: Tuszyński, George
: APPLICANT: Williams, Taft
: TITLE OF INVENTION: RETROINVERSO POLYPEPTIDES THAT MIMIC OR INHIBIT
: FILE REFERENCE: 07206-0021
: CURRENT APPLICATION NUMBER: US/09/197,770B
: CURRENT FILING DATE: 1998-11-23
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 17
: LENGTH: 9
: TYPE: PRP
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: synthetic
: OTHER INFORMATION: fragment/ analog of thrombospondin
: US-09-197-770B-17

Query Match
Best Local Similarity 94.9%; Score 56; DB 5; Length 9;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WSPCSVTGC 9
   |:|||||
Db 1 WTPCSVTGC 9

RESULT 8
: US-09-978-189-301
: Sequence 301, Application US/09978189
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gertlisen, Mary E.
: APPLICANT: Goddard, Audrey
```

```

: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavits, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paonli, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630P1C7
: CURRENT APPLICATION NUMBER: US/09/978,189
: CURRENT FILING DATE: 2001-10-15
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: PRIOR APPLICATION NUMBER: 60/078004
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: 60/078886
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078936
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078939
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079664
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079663
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079786
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079920
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/079923
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/080105
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080107
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080165
```

PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080194  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080327  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080328  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080333  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080334  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/081070  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081049  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081071  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081195  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081955  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081819  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081952  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081838  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082568  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082569  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082704  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082804  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082700  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082797  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082796  
 PRIOR FILING DATE: 1998-04-23  
 PRIOR APPLICATION NUMBER: 60/083336  
 PRIOR FILING DATE: 1998-04-27  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083392  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083495  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083496  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083499  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083554  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083558  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083559  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083500  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083742  
 PRIOR FILING DATE: 1998-04-30  
 PRIOR APPLICATION NUMBER: 60/084366  
 PRIOR FILING DATE: 1998-05-05

PRIOR APPLICATION NUMBER: 60/084414  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084441  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084639  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084640  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084598  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084643  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085338  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085323  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085582  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085700  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085689  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085579  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085580  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085573  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085704  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085697  
 PRIOR FILING DATE: 1998-05-15

Query Match 94.9%; Score 56; DB 5; Length 525;  
 Best Local Similarity 88.9%; Pred. No. 0.093;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSTG 9  
 |||||  
 Db 445 WSPCSTG 453

RESULT 9  
 US-09-763-397a-8.rapn  
 ; Sequence 301, Application US/09978192  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Bolstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kijavlin, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.

APPLICANT: Pan, James;  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630PIC9  
CURRENT APPLICATION NUMBER: US/09/978.192  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01

PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639

PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697  
PRIOR FILING DATE: 1998-05-15

Query Match 94.9%; Score 56; DB 5; Length 525;  
Best Local Similarity 88.9%; Pred. No. 0.093;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSPCSTCG 9  
|||||  
DB 445 WSPCTVTCG 453

RESULT 10  
US-09-978-697-301  
: Sequence 301. Application US/09978697  
: GENERAL INFORMATION:  
: APPLICANT: Ashkenazi, Avi  
: APPLICANT: Baker Kevin P.  
: APPLICANT: Botsstein, David  
: APPLICANT: Desnoyers, Luc  
: APPLICANT: Eaton, Dan  
: APPLICANT: Ferrara, Napoleon  
: APPLICANT: Filvaroff, Ellen  
: APPLICANT: Fong, Sherman  
: APPLICANT: Gao, Wei-Qiang  
: APPLICANT: Gerber, Hanspeter  
: APPLICANT: Gerritsen, Mary E.  
: APPLICANT: Goddard, Audrey  
: APPLICANT: Godowski, Paul J.  
: APPLICANT: Grimaldi, J. Christopher  
: APPLICANT: Gurney, Austin L.  
: APPLICANT: Hillan, Kenneth J.  
: APPLICANT: Kijavlin, Ivar J.  
: APPLICANT: Kuo, Sophia S.  
: APPLICANT: Napier, Mary A.  
: APPLICANT: Pan, James;  
: APPLICANT: Paoni, Nicholas F.  
: APPLICANT: Roy, Margaret Ann  
: APPLICANT: Shelton, David L.  
: APPLICANT: Stewart, Timothy A.  
: APPLICANT: Tumas, Daniel  
: APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2630PIC27  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049



;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081071  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081955  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081819  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081952  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082568  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082804  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082700  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082797  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082796  
;; PRIOR FILING DATE: 1998-04-23  
;; PRIOR APPLICATION NUMBER: 60/083336  
;; PRIOR FILING DATE: 1998-04-27  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083392  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/083366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-5-07

;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697  
;; PRIOR FILING DATE: 1998-05-15

Query Match 94.9%; Score 56; DB 5; Length 525;  
Best Local Similarity 88.9%; Pred. No. 0.093;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WSPCSTVC 9  
||||:||||  
Db 445 WSPCSTVC 453

RESULT 11  
US-09-978-824-301  
;; Sequence 301, Application US/09978824  
;; GENERAL INFORMATION:  
;; APPLICANT: Ashkenazi, Avi  
;; APPLICANT: Baker Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Eaton, Dan  
;; APPLICANT: Ferrara, Napoleon  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Guirney, Austin L.  
;; APPLICANT: Hillan, Kenneth J.  
;; APPLICANT: Kljavin, Ivar J.  
;; APPLICANT: Kuo, Sophia S.  
;; APPLICANT: Napier, Mary A.  
;; APPLICANT: Pan, James;  
;; APPLICANT: Paonli, Nicholas F.  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Shelton, David L.  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Williams, P. Mickey  
;; APPLICANT: Wood, William I.  
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
;; FILE REFERENCE: P2650PIC14  
;; CURRENT FILING DATE: 2001-10-17  
;; PRIOR APPLICATION NUMBER: 09/918585

1	PRIOR FILING DATE:	2001-07-30
2	PRIOR APPLICATION NUMBER:	60/062250
3	PRIOR FILING DATE:	1997-10-17
4	PRIOR APPLICATION NUMBER:	60/064249
5	PRIOR FILING DATE:	1997-11-03
6	PRIOR APPLICATION NUMBER:	60/065311
7	PRIOR FILING DATE:	1997-11-13
8	PRIOR APPLICATION NUMBER:	60/066364
9	PRIOR FILING DATE:	1997-11-21
10	PRIOR APPLICATION NUMBER:	60/077450
11	PRIOR FILING DATE:	1998-03-10
12	PRIOR APPLICATION NUMBER:	60/077652
13	PRIOR FILING DATE:	1998-03-11
14	PRIOR APPLICATION NUMBER:	60/077641
15	PRIOR FILING DATE:	1998-03-11
16	PRIOR APPLICATION NUMBER:	60/077649
17	PRIOR FILING DATE:	1998-03-11
18	PRIOR APPLICATION NUMBER:	60/077791
19	PRIOR FILING DATE:	1998-03-12
20	PRIOR APPLICATION NUMBER:	60/078004
21	PRIOR FILING DATE:	1998-03-13
22	PRIOR APPLICATION NUMBER:	60/078886
23	PRIOR FILING DATE:	1998-03-20
24	PRIOR APPLICATION NUMBER:	60/078936
25	PRIOR FILING DATE:	1998-03-20
26	PRIOR APPLICATION NUMBER:	60/078910
27	PRIOR FILING DATE:	1998-03-20
28	PRIOR APPLICATION NUMBER:	60/078939
29	PRIOR FILING DATE:	1998-03-20
30	PRIOR APPLICATION NUMBER:	60/079294
31	PRIOR FILING DATE:	1998-03-25
32	PRIOR APPLICATION NUMBER:	60/079656
33	PRIOR FILING DATE:	1998-03-26
34	PRIOR APPLICATION NUMBER:	60/079664
35	PRIOR FILING DATE:	1998-03-27
36	PRIOR APPLICATION NUMBER:	60/079689
37	PRIOR FILING DATE:	1998-03-27
38	PRIOR APPLICATION NUMBER:	60/079663
39	PRIOR FILING DATE:	1998-03-27
40	PRIOR APPLICATION NUMBER:	60/079728
41	PRIOR FILING DATE:	1998-03-27
42	PRIOR APPLICATION NUMBER:	60/079786
43	PRIOR FILING DATE:	1998-03-27
44	PRIOR APPLICATION NUMBER:	60/079920
45	PRIOR FILING DATE:	1998-03-30
46	PRIOR APPLICATION NUMBER:	60/079923
47	PRIOR FILING DATE:	1998-03-30
48	PRIOR APPLICATION NUMBER:	60/080105
49	PRIOR FILING DATE:	1998-03-31
50	PRIOR APPLICATION NUMBER:	60/080107
51	PRIOR FILING DATE:	1998-03-31
52	PRIOR APPLICATION NUMBER:	60/080165
53	PRIOR FILING DATE:	1998-03-31
54	PRIOR APPLICATION NUMBER:	60/080194
55	PRIOR FILING DATE:	1998-03-31
56	PRIOR APPLICATION NUMBER:	60/080327
57	PRIOR FILING DATE:	1998-04-01
58	PRIOR APPLICATION NUMBER:	60/080328
59	PRIOR FILING DATE:	1998-04-01
60	PRIOR APPLICATION NUMBER:	60/080333
61	PRIOR FILING DATE:	1998-04-01
62	PRIOR APPLICATION NUMBER:	60/080334
63	PRIOR FILING DATE:	1998-04-01
64	PRIOR APPLICATION NUMBER:	60/081070
65	PRIOR FILING DATE:	1998-04-08
66	PRIOR APPLICATION NUMBER:	60/081049
67	PRIOR FILING DATE:	1998-04-08
68	PRIOR APPLICATION NUMBER:	60/081071
69	PRIOR FILING DATE:	1998-04-08
70	PRIOR APPLICATION NUMBER:	60/081195
71	PRIOR FILING DATE:	1998-04-08
72	PRIOR APPLICATION NUMBER:	60/081203
73	PRIOR FILING DATE:	1998-04-09

1	PRIOR APPLICATION NUMBER: 60/081229
2	PRIOR FILING DATE: 1998-04-09
3	PRIOR APPLICATION NUMBER: 60/081955
4	PRIOR FILING DATE: 1998-04-15
5	PRIOR APPLICATION NUMBER: 60/081817
6	PRIOR FILING DATE: 1998-04-15
7	PRIOR APPLICATION NUMBER: 60/081819
8	PRIOR FILING DATE: 1998-04-15
9	PRIOR APPLICATION NUMBER: 60/081952
10	PRIOR FILING DATE: 1998-04-15
11	PRIOR APPLICATION NUMBER: 60/081838
12	PRIOR FILING DATE: 1998-04-15
13	PRIOR APPLICATION NUMBER: 60/082568
14	PRIOR FILING DATE: 1998-04-21
15	PRIOR APPLICATION NUMBER: 60/082569
16	PRIOR FILING DATE: 1998-04-21
17	PRIOR APPLICATION NUMBER: 60/082704
18	PRIOR FILING DATE: 1998-04-22
19	PRIOR APPLICATION NUMBER: 60/082804
20	PRIOR FILING DATE: 1998-04-22
21	PRIOR APPLICATION NUMBER: 60/082700
22	PRIOR FILING DATE: 1998-04-22
23	PRIOR APPLICATION NUMBER: 60/082797
24	PRIOR FILING DATE: 1998-04-22
25	PRIOR APPLICATION NUMBER: 60/082796
26	PRIOR FILING DATE: 1998-04-23
27	PRIOR APPLICATION NUMBER: 60/083336
28	PRIOR FILING DATE: 1998-04-27
29	PRIOR APPLICATION NUMBER: 60/083322
30	PRIOR FILING DATE: 1998-04-28
31	PRIOR APPLICATION NUMBER: 60/083392
32	PRIOR FILING DATE: 1998-04-29
33	PRIOR APPLICATION NUMBER: 60/083495
34	PRIOR FILING DATE: 1998-04-29
35	PRIOR APPLICATION NUMBER: 60/083496
36	PRIOR FILING DATE: 1998-04-29
37	PRIOR APPLICATION NUMBER: 60/083499
38	PRIOR FILING DATE: 1998-04-29
39	PRIOR APPLICATION NUMBER: 60/083545
40	PRIOR FILING DATE: 1998-04-29
41	PRIOR APPLICATION NUMBER: 60/083554
42	PRIOR FILING DATE: 1998-04-29
43	PRIOR APPLICATION NUMBER: 60/083558
44	PRIOR FILING DATE: 1998-04-29
45	PRIOR APPLICATION NUMBER: 60/083559
46	PRIOR FILING DATE: 1998-04-29
47	PRIOR APPLICATION NUMBER: 60/083530
48	PRIOR FILING DATE: 1998-04-29
49	PRIOR APPLICATION NUMBER: 60/083742
50	PRIOR FILING DATE: 1998-04-30
51	PRIOR APPLICATION NUMBER: 60/084366
52	PRIOR FILING DATE: 1998-05-05
53	PRIOR APPLICATION NUMBER: 60/084414
54	PRIOR FILING DATE: 1998-05-06
55	PRIOR APPLICATION NUMBER: 60/084441
56	PRIOR FILING DATE: 1998-05-06
57	PRIOR APPLICATION NUMBER: 60/084637
58	PRIOR FILING DATE: 1998-05-07
59	PRIOR APPLICATION NUMBER: 60/084639
60	PRIOR FILING DATE: 1998-05-07
61	PRIOR APPLICATION NUMBER: 60/084640
62	PRIOR FILING DATE: 1998-05-07
63	PRIOR APPLICATION NUMBER: 60/084598
64	PRIOR FILING DATE: 1998-05-07
65	PRIOR APPLICATION NUMBER: 60/084600
66	PRIOR FILING DATE: 1998-05-07
67	PRIOR APPLICATION NUMBER: 60/084627
68	PRIOR FILING DATE: 1998-05-07
69	PRIOR APPLICATION NUMBER: 60/084643
70	PRIOR FILING DATE: 1998-05-07
71	PRIOR APPLICATION NUMBER: 60/085339
72	PRIOR FILING DATE: 1998-05-13
73	PRIOR APPLICATION NUMBER: 60/085338

```
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
```

```
Query Match          94.9%; Score 56; DB 5; Length 525;
Best Local Similarity 88.9%; Pred. No. 0.093;
Matches      8; Conservative      1; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      1 WSPCSVTGC 9
        ||||:||||
Db      445 WSPCSVTGC 453
```

```
RESULT 12
US-09-970-966-207
; Sequence 207, Application US/09970966
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesn, David Alan
; APPLICANT: Filing, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970.966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 207
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-966-207
```

```
Query Match          91.5%; Score 54; DB 5; Length 787;
Best Local Similarity 77.8%; Pred. No. 0.25;
Matches      7; Conservative      2; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      1 WSPCSVTGC 9
        ||||:||||
Db      490 WSPCSVTGC 498
```

```
RESULT 13
US-09-970-966-186
; Sequence 186, Application US/09970966
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesn, David Alan
; APPLICANT: Filing, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970.966
; CURRENT FILING DATE: 2001-10-02
```

```
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-966-186
```

```
Query Match          91.5%; Score 54; DB 5; Length 807;
Best Local Similarity 77.8%; Pred. No. 0.25;
Matches      7; Conservative      2; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      1 WSPCSVTGC 9
        ||||:||||
Db      510 WSPCSVTGC 518
```

```
RESULT 14
US-09-611-526-3833
; Sequence 3833, Application US/09611526
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: KANAI, YURI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: KOJIMA, SHINICHI
; APPLICANT: OTSUKI, TERSUJI
; APPLICANT: KOGA, HISASHI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH CDNAS
; TITLE OF INVENTION: AND THEIR USES
; FILE REFERENCE: 08335/0122
; CURRENT APPLICATION NUMBER: US/09/611.526
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: JP 1999-194486
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: JP 2000-118774
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183765
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 4484
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3833
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-611-526-3833
```

```
Query Match          91.5%; Score 54; DB 5; Length 807;
Best Local Similarity 77.8%; Pred. No. 0.25;
Matches      7; Conservative      2; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      1 WSPCSVTGC 9
        ||||:||||
Db      510 WSPCSVTGC 518
```

```
RESULT 15
US-09-197-770B-14
; Sequence 14, Application US/09197770B
; GENERAL INFORMATION:
; APPLICANT: Tuszyński, George
; APPLICANT: Williams, Taffy
; APPLICANT: Actor, Paul
; TITLE OF INVENTION: RETROINVERSO POLYPEPTIDES THAT MIMIC OR INHIBIT
; TITLE OF INVENTION: THROMBOSPONDIN ACTIVITY
; FILE REFERENCE: 07206-0021
```

US-09-197-770B-14  
: CURRENT APPLICATION NUMBER: US/09/197,770B  
: CURRENT FILING DATE: 1998-11-23  
: NUMBER OF SEQ ID NOS: 38  
: SOFTWARE: Patentl Ver. 2.0  
: SEQ ID NO 14  
: LENGTH: 9  
: TYPE: PRT  
: ORGANISM: Artificial Sequence  
: FEATURE:  
: OTHER INFORMATION: Description of Artificial Sequence: synthetic  
: OTHER INFORMATION: fragment/ analog of thrombospondin

Query Match 89.88; Score 53; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WSPCSVTC 8  
Db 1 WSPCSVTC 8

Search completed: January 29, 2002, 10:58:11  
Job time: 2507 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:01 ; Search time 1760.55 Seconds  
(without alignments)  
3.312 Million cell updates/sec

Title: US-09-763-397A-7  
Perfect score: 109  
Sequence: 1 DIEKRICKMEKCSVFNNVNS 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 segs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main: \*  
1: /cgn2\_6/ptodata/2/paa/PCRTUS\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep: \*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep: \*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep: \*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep: \*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep: \*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep: \*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep: \*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep: \*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep: \*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep: \*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep: \*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep: \*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep: \*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep: \*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep: \*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep: \*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep: \*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep: \*  
24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	100.0	21	US-09-763-397A-7	Sequence 7, Appl1
2	109	100.0	350	US-09-763-397A-2	Sequence 2, Appl1
3	109	100.0	396	US-09-820-843A-31	Sequence 31, Appl1
4	109	100.0	412	US-07-677-539B-6	Sequence 6, Appl1
5	109	100.0	412	US-07-727-636-6	Sequence 6, Appl1
6	109	100.0	412	US-08-948-885-18	Sequence 18, Appl1
7	109	100.0	423	US-08-932-929A-1	Sequence 1, Appl1
8	109	100.0	424	US-07-842-694-2	Sequence 2, Appl1
9	109	100.0	424	US-08-760-797-1	Sequence 1, Appl1

10	109	100.0	424	US-08-760-797-3	Sequence 3, Appl1
11	109	100.0	424	US-08-903-084-1	Sequence 1, Appl1
12	109	100.0	424	US-08-903-084-3	Sequence 3, Appl1
13	109	100.0	424	US-08-932-929-1	Sequence 1, Appl1
14	109	100.0	424	US-08-932-929-3	Sequence 3, Appl1
15	109	100.0	424	US-08-932-929A-3	Sequence 3, Appl1
16	107	98.2	388	US-07-677-539B-5	Sequence 5, Appl1
17	107	98.2	388	US-07-727-636-5	Sequence 5, Appl1
18	91	83.5	21	PCT-US00-24802-3875	Sequence 3875, Ap
19	91	83.5	21	PCT-US00-33549-68	Sequence 68, Appl
20	91	83.5	21	PCT-US00-34318-27	Sequence 27, Appl
21	91	83.5	21	PCT-US00-35516-12	Sequence 12, Appl
22	91	83.5	21	PCT-US94-04832A-48	Sequence 48, Appl
23	91	83.5	21	PCT-US99-13923-41	Sequence 41, Appl
24	91	83.5	21	PCT-US99-13929-78	Sequence 78, Appl
25	91	83.5	21	US-08-121-101-17	Sequence 17, Appl
26	91	83.5	21	US-08-197-484-97	Sequence 97, Appl
27	91	83.5	21	US-08-205-713C-1042	Sequence 1042, Ap
28	91	83.5	21	US-08-229-275-48	Sequence 48, Appl
29	91	83.5	21	US-08-328-912B-32	Sequence 32, Appl
30	91	83.5	21	US-08-464-496-17	Sequence 17, Appl
31	91	83.5	21	US-08-488-320A-48	Sequence 48, Appl
32	91	83.5	21	US-08-718-490A-32	Sequence 32, Appl
33	91	83.5	21	US-08-738-942-93	Sequence 93, Appl
34	91	83.5	21	US-08-788-822A-12	Sequence 12, Appl
35	91	83.5	21	US-08-815-396-34	Sequence 34, Appl
36	91	83.5	21	US-08-820-360-97	Sequence 97, Appl
37	91	83.5	21	US-08-926-296-20	Sequence 20, Appl
38	91	83.5	21	US-09-100-409-54	Sequence 54, Appl
39	91	83.5	21	US-09-100-415-41	Sequence 41, Appl
40	91	83.5	21	US-09-165-848-13	Sequence 13, Appl
41	91	83.5	21	US-09-203-143-7	Sequence 7, Appl1
42	91	83.5	21	US-09-226-775-2	Sequence 2, Appl1
43	91	83.5	21	US-09-239-043-2573	Sequence 2573, Ap
44	91	83.5	21	US-09-260-714A-2	Sequence 2, Appl1
45	91	83.5	21	US-09-260-714A-2	Sequence 2, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-763-397A-7  
; Sequence 7, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; SECRETARY OF THE Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-763-397A-7

Query Match 100.0%; Score 109; DB 21; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.2e-09;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
By 1 DIEKRICKMEKCSVFNNVNS 21

Db 1 DIEKICKMEKCSSVFNVNS 21  
|||||

RESULT 2  
US-09-763-397A-2  
; Sequence 2, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci  
; FILE REFERENCE: 6395-57049  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR FILING DATE: 2001-02-16  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Recombinant DNA/Protein  
US-09-763-397A-2

Query Match 100.0%; Score 109; DB 21; Length 350;  
Best Local Similarity 100.0%; Pred. No. 6,6e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DIEKICKMEKCSSVFNVNS 21  
|||||

Db 61 DIEKICKMEKCSSVFNVNS 81

RESULT 3  
US-09-820-843A-31  
; Sequence 31, Application US/09820843A  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEI  
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES  
; FILE REFERENCE: 063915  
; CURRENT APPLICATION NUMBER: US/09/820,843A  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 31  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Circumsporozoite (CS) protein  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: g114493889  
US-09-820-843A-31

Query Match 100.0%; Score 109; DB 22; Length 396;  
Best Local Similarity 100.0%; Pred. No. 7,5e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DIEKICKMEKCSSVFNVNS 21  
|||||

Db 362 DIEKICKMEKCSSVFNVNS 382

RESULT 4  
US-07-677-539B-6  
; Sequence 6, Application US/07677539B  
; GENERAL INFORMATION:  
; APPLICANT: LAL, ALTA F.  
; APPLICANT: GOLDMAN, IRA F.  
; TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM  
; TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L STREET, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/677,539B  
FILING DATE: 19911205  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: SCOTT, WATSON T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: 5683/91540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3067  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 412 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-677-539B-6

Query Match 100.0%; Score 109; DB 3; Length 412;  
Best Local Similarity 100.0%; Pred. No. 7,8e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DIEKICKMEKCSSVFNVNS 21  
|||||

Db 378 DIEKICKMEKCSSVFNVNS 398

RESULT 5  
US-07-727-636-6  
; Sequence 6, Application US/07727636  
; GENERAL INFORMATION:  
; APPLICANT: LAL, ALTA F.  
; APPLICANT: GOLDMAN, IRA F.  
; TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM  
; TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L STREET, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/727,636  
FILING DATE: 19910710  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: SCOTT, WATSON T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: 5663/91540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3067  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 412 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-727-636-6

Query Match 100.0%; Score 109; DB 3; Length 412;  
Best Local Similarity 100.0%; Pred. No. 7.8e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKICKMEKCSVFNVNVS 21  
|||||  
DB 378 DIEKICKMEKCSVFNVNVS 398

RESULT 6  
US-08-948-885-18  
Sequence 18, Application US/08948885  
GENERAL INFORMATION:  
APPLICANT: Jessell, Thomas M. and Avihu Klar  
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A  
TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/948,885  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/313,288  
FILING DATE: January 5, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0526  
TELEX:  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 412 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-948-885-18

Query Match 100.0%; Score 109; DB 13; Length 412;  
Best Local Similarity 100.0%; Pred. No. 7.8e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKICKMEKCSVFNVNVS 21  
|||||  
DB 378 DIEKICKMEKCSVFNVNVS 398

RESULT 7  
US-08-932-929A-1  
Sequence 1, Application US/08932929A  
GENERAL INFORMATION:  
APPLICANT: De Wilde, Michel  
TITLE OF INVENTION: Hybrid Protein Between CS  
TITLE OF INVENTION: from Plasmidium and HbsAG  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/932,929A  
FILING DATE: 18-SEPT-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/663,371  
FILING DATE: 13-JUNE-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: B45015-1FWC2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 423 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-932-929A-1

Query Match 100.0%; Score 109; DB 13; Length 423;  
Best Local Similarity 100.0%; Pred. No. 8.1e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKICKMEKCSVFNVNVS 21  
|||||  
DB 172 DIEKICKMEKCSVFNVNVS 192

RESULT 8  
US-07-842-694-2  
Sequence 2, Application US/07842694  
GENERAL INFORMATION:  
APPLICANT: Cohen, Joseph





TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-760-797-3

Query Match 100.0%; Score 109; DB 11; Length 424;  
Best Local Similarity 100.0%; Pred. No. 8.1e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKICKMEKCSSVFNVNS 21  
|||||

DB 169 DIEKICKMEKCSSVFNVNS 189

## RESULT 11

US-08-903-084-1  
Sequence 1, Application US/08903084  
GENERAL INFORMATION:  
APPLICANT: Bruck, Claudine  
APPLICANT: Francotte, Myriam  
APPLICANT: Kummerl, Suzanne  
APPLICANT: Siaoui, Moncef  
APPLICANT: Wijndale, Frans  
TITLE OF INVENTION: VACCINE FORMULATIONS CONTAINING  
TITLE OF INVENTION: 3-O-DEACYLATED MONOPHOSPHORYL LIPID A AND A CARRIER  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19046  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/903,084  
FILING DATE: 17-JUL-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/741,575  
FILING DATE: 30-OCT-1996  
APPLICATION NUMBER: 08/303,542  
FILING DATE: 09-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Kerekes, Zoltan  
REGISTRATION NUMBER: 38,938  
REFERENCE/DOCKET NUMBER: B45101  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5024  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-903-084-1

Query Match 100.0%; Score 109; DB 13; Length 424;  
Best Local Similarity 100.0%; Pred. No. 8.1e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKICKMEKCSSVFNVNS 21  
|||||

DB 169 DIEKICKMEKCSSVFNVNS 189

RESULT 12  
US-08-903-084-3

Sequence 3, Application US/08903084  
GENERAL INFORMATION:  
APPLICANT: Bruck, Claudine  
APPLICANT: Francotte, Myriam  
APPLICANT: Kummerl, Suzanne  
APPLICANT: Siaoui, Moncef  
APPLICANT: Wijndale, Frans  
TITLE OF INVENTION: VACCINE FORMULATIONS CONTAINING  
TITLE OF INVENTION: 3-O-DEACYLATED MONOPHOSPHORYL LIPID A AND A CARRIER  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19046  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/903,084  
FILING DATE: 17-JUL-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/741,575  
FILING DATE: 30-OCT-1996  
APPLICATION NUMBER: 08/303,542  
FILING DATE: 09-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Kerekes, Zoltan  
REGISTRATION NUMBER: 38,938  
REFERENCE/DOCKET NUMBER: B45101  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5024  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-903-084-3

Query Match 100.0%; Score 109; DB 13; Length 424;  
Best Local Similarity 100.0%; Pred. No. 8.1e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKICKMEKCSSVFNVNS 21  
|||||

DB 169 DIEKICKMEKCSSVFNVNS 189

RESULT 13  
US-08-932-929-1

Sequence 1, Application US/08932929  
GENERAL INFORMATION:  
APPLICANT: De Wilde, Michel  
APPLICANT: Cohen, Joseph  
TITLE OF INVENTION: HYBRID PROTEIN BETWEEN CS FROM  
TITLE OF INVENTION: PLASMODIUM AND HBSAG  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corp./Corporate  
ADDRESSEE: Intellectual Property

STREET: P.O. Box 1539 - UM2220  
CITY: King of Prussia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/932.929  
FILING DATE: 18-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/663.371  
FILING DATE:  
APPLICATION NUMBER: US/08/244.085  
FILING DATE: 21-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: B45015-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610 270-5096  
TELEFAX: 610 270-5090  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-932-929-1

Query Match 100.0%; Score 109; DB 13; Length 424;  
Best Local Similarity 100.0%; Pred. No. 8.1e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKICKMEKCSVPNVNS 21  
|||||

DB 169 DIEKICKMEKCSVPNVNS 189

RESULT 14  
US-08-932-929-3  
Sequence 3, Application US/08932929  
GENERAL INFORMATION:  
APPLICANT: De Wilde, Michel  
APPLICANT: Cohen, Joseph  
TITLE OF INVENTION: HYBRID PROTEIN BETWEEN CS FROM  
TITLE OF INVENTION: PLASMODIUM AND HBSAG  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corp./Corporate  
ADDRESSEE: Intellectual Property  
STREET: P.O. Box 1539 - UM2220  
CITY: King of Prussia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/932.929  
FILING DATE: 18-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/663.371

FILING DATE:  
APPLICATION NUMBER: US 08/244.085  
FILING DATE: 21-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: B45015-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610 270-5096  
TELEFAX: 610 270-5090  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-932-929-3

Query Match 100.0%; Score 109; DB 13; Length 424;  
Best Local Similarity 100.0%; Pred. No. 8.1e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIEKICKMEKCSVPNVNS 21  
|||||

DB 169 DIEKICKMEKCSVPNVNS 189

RESULT 15  
US-08-932-929A-3  
Sequence 3, Application US/08932929A  
GENERAL INFORMATION:  
APPLICANT: De Wilde, Michel  
APPLICANT: Cohen, Joseph  
TITLE OF INVENTION: Hybrid Protein Between CS  
TITLE OF INVENTION: from Plasmodium and HBSAG  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/932.929A  
FILING DATE: 18-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/663.371  
FILING DATE: 13-JUNE-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: B45015-1FWC2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-932-929A-3

Query Match 100.0%; Score 109; DB 13; Length 424;  
Best Local Similarity 100.0%; Pred. No. 8.1e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 DIEKKICKMEKCSYFNNVNS 21  
|||||  
Db 169 DIEKKICKMEKCSYFNNVNS 189

Search completed: January 29, 2002, 10:56:02  
Job time: 2398 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:09 : Search time 120.95 Seconds  
(without alignments)  
6.354 Million cell updates/sec

Title: US-09-763-397A-7

Perfect score: 109

Sequence: 1 DIEKRICMEKCSVFNVNS 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCY\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US50\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.4	167	5	US-09-611-526-3845	Sequence 3845, Ap
2	43	39.4	115	US-09-708-427-35760	Sequence 35760, A
3	43	39.4	117	US-09-708-427-35758	Sequence 35758, A
4	42.5	39.0	58	US-09-620-111B-3030	Sequence 3030, Ap
5	42.5	39.0	68	US-09-620-111B-3029	Sequence 3029, Ap
6	42	38.5	15	US-09-894-018-273	Sequence 3029, Ap
7	42	38.5	276	US-09-894-018-141	Sequence 273, Ap
8	42	38.5	372	US-09-708-427-22442	Sequence 141, Ap
9	42	38.5	374	US-09-708-427-22442	Sequence 141, Ap
10	42	38.5	456	US-09-620-394B-5663	Sequence 22442, A
11	41	37.6	150	US-09-894-018-121	Sequence 5663, Ap
12	41	37.6	238	US-09-708-427-39156	Sequence 121, App
13	41	37.6	239	US-09-708-427-26073	Sequence 39156, A
14	41	37.6	241	US-09-605-703B-292	Sequence 26073, A
15	41	37.6	256	US-09-708-427-4856	Sequence 292, App
16	41	37.6	312	US-09-708-427-4855	Sequence 4856, App
17	41	37.6	638	US-09-605-703B-290	Sequence 4855, Ap
18	41	37.6	681	US-09-708-427-23777	Sequence 290, App
19	41	37.6	707	US-09-708-427-23776	Sequence 23777, A
20	40	36.7	241	US-09-708-427-23775	Sequence 23776, A
21	40	36.7	690	US-09-708-427-25306	Sequence 23775, A
22	40	36.7	691	US-09-708-427-8327	Sequence 25306, A
23	40	36.7	860	US-09-708-427-8326	Sequence 8327, Ap
24	40	36.7	923	US-09-708-427-8325	Sequence 8326, Ap
25	40	36.7	956	US-09-708-427-23705	Sequence 8325, Ap
26	40	36.7	1006	US-09-708-427-23704	Sequence 23705, A
				US-09-708-427-23703	Sequence 23704, A
					Sequence 23703, A

27	40	36.7	1221	5	US-09-919-891-2	Sequence 2, Appl1
28	39.5	36.2	317	5	US-09-985-153-54	Sequence 54, Appl
29	39.5	36.2	317	6	US-10-006-867-116	Sequence 116, App
30	39.5	36.2	366	5	US-09-708-427-33001	Sequence 33001, A
31	39.5	36.2	382	5	US-09-708-427-33000	Sequence 33000, A
32	39.5	36.2	464	5	US-09-708-427-32999	Sequence 32999, A
33	39	35.8	35	7	US-60-323-991-23	Sequence 23, Appl
34	39	35.8	51	5	US-09-708-427-64932	Sequence 64932, A
35	39	35.8	120	5	US-09-760-446A-2210	Sequence 2210, Ap
36	39	35.8	270	5	US-09-620-394B-1049	Sequence 1049, Ap
37	39	35.8	272	5	US-09-620-394B-1048	Sequence 1048, Ap
38	39	35.8	272	5	US-09-708-427-24404	Sequence 24404, A
39	39	35.8	324	5	US-09-708-427-24403	Sequence 24403, A
40	39	35.8	336	5	US-09-620-394B-1047	Sequence 1047, Ap
41	39	35.8	341	5	US-09-708-427-9857	Sequence 9857, Ap
42	39	35.8	366	5	US-09-708-427-9856	Sequence 9856, Ap
43	39	35.8	370	5	US-09-708-427-21983	Sequence 21983, A
44	39	35.8	372	5	US-09-708-427-21982	Sequence 21982, A
45	39	35.8	379	5	US-09-708-427-21981	Sequence 21981, A

## ALIGNMENTS

```
RESULT 1
US-09-611-526-3845
; Sequence 3845, Application US/09611526
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: KAWAI, YURI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: KOJIMA, SHINICHI
; APPLICANT: OTSUKI, TETSUHI
; APPLICANT: KOGA, HISASHI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH CDNAS
; FILE REFERENCE: 08335/0122
; CURRENT APPLICATION NUMBER: US/09/611,526
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: JP 1999-194486
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: JP 2000-118774
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183765
; NUMBER OF SEQ ID NOS: 4484
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3845
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-611-526-3845

Query Match      40.4%; Score 44; DB 5; Length 167;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      4 KRICKMEKCSVFN 17
DB      147 EKSKYKCKGEAFN 160

RESULT 2
US-09-708-427-35760
; Sequence 35760, Application US/09708427
; GENERAL INFORMATION:
```

RESULT 4  
US-09-620-111B-3030  
; Sequence 3030, Application US/09620111B  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; TITLE OF INVENTION: Thereby

```

RESULT 6
US-09-894-018-273
; Sequence 273, Application US/09894018
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Dennis
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING

```

```
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-2003.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 273
; LENGTH: 15
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-894-018-273
```

```
Query Match          38.5%; Score 42; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      13 SSVFNVNS 21
        |||||
Db       1 SSVFNVNS 9
```

```
RESULT 7
US-09-894-018-141
; Sequence 141, Application US/09894018
; GENERAL INFORMATION:
; APPLICANT: PRIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; FILE REFERENCE: 39963-2003.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 141
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PfHFL
US-09-894-018-141
```

```
Query Match          38.5%; Score 42; DB 5; Length 276;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      13 SSVFNVNS 21
        |||||
Db       102 SSVFNVNS 110
```

```
RESULT 8
US-09-708-427-22442
; Sequence 22442, Application US/09708427
```

```
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22442
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..372
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..372
; OTHER INFORMATION: Ceres Seq. ID 1840904
US-09-708-427-22442
```

```
Query Match          38.5%; Score 42; DB 5; Length 372;
Best Local Similarity 57.1%; Pred. No. 55;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
OY      4 KKICKMEKSSVFN 17
        | | | | | | |
Db       234 KPITKLEDCESFPN 247
```

```
RESULT 9
US-09-620-394B-5663
; Sequence 5663, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypeptid
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 5663
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..374
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..374
; OTHER INFORMATION: Ceres Seq. ID 1394837
US-09-620-394B-5663
```

```
Query Match          38.5%; Score 42; DB 5; Length 374;
Best Local Similarity 57.1%; Pred. No. 55;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
OY      4 KKICKMEKSSVFN 17
        | | | | | | |
Db       234 KPITKLEDCESFPN 247
```

```
RESULT 10
US-09-894-018-121
; Sequence 121, Application US/09894018
; GENERAL INFORMATION:
; APPLICANT: PRIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
```

```
APPLICANT: Chestnut, Robert
APPLICANT: Livingston, Brian
APPLICANT: Baker, Denlew
APPLICANT: Newman, Mark
APPLICANT: Brown, David
TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
FILE REFERENCE: 39963-20033.00
CURRENT APPLICATION NUMBER: US/09/894,018
CURRENT FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: PCT/US00/35568
PRIOR FILING DATE: 2000-12-28
PRIOR APPLICATION NUMBER: US 60/173,390
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: US 60/284,221
PRIOR FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 368
SOFTWARE: PastedSeq for Windows Version 4.0
SEQ ID NO 121
LENGTH: 456
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PICTL/HTL(N)
US-09-894-018-121
```

```
Query Match          38.5%; Score 42; DB 5; Length 456;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 13 SSVFNVNS 21
||| |||||
DB 21 SSVFNVNS 29
```

```
RESULT 11
US-09-708-427-39156
; Sequence 39156, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39156
; LENGTH: 150
; TYPE: PRT
; ORGANISM: zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..150
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..150
; OTHER INFORMATION: Ceres Seq. ID 1852376
US-09-708-427-39156
```

```
Query Match          37.6%; Score 41; DB 5; Length 150;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 1 DIEKKICKMEKCS 14
||| ||| : :|
DB 84 DLKKFKCKFGRYTS 97
```

```
RESULT 12
US-09-708-427-26073
```

```
; Sequence 26073, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26073
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..238
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..238
; OTHER INFORMATION: Ceres Seq. ID 1815731
US-09-708-427-26073
```

```
Query Match          37.6%; Score 41; DB 5; Length 238;
Best Local Similarity 45.0%; Pred. No. 50;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
OY 2 IEKKICKMEKCSSVFNVNS 21
||| : || : ||| |
DB 118 IEKQHKLEAKSIFNEAES 137
```

```
RESULT 13
US-09-605-703B-292
; Sequence 292, Application US/09605703B
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-1296CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 292
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-292
```

```
Query Match          37.6%; Score 41; DB 5; Length 239;
Best Local Similarity 33.3%; Pred. No. 51;
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 DIEKKICKMEKCSSV 15
|| : || : ||| :|
DB 210 NVETRHCKRRCAL 224
```

```
RESULT 14
US-09-708-427-4856
; Sequence 4856, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
```



```

: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
: TITLE OF INVENTION: THEREBY
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4856
: LENGTH: 241
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc-feature
: LOCATION: 1..241
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc-feature
: LOCATION: 1..241
: OTHER INFORMATION: Ceres Seq. ID 1813257
US-09-708-427-4856

```

```

Query Match          37.6%; Score 41; DB 5; Length 241;
Best Local Similarity 38.9%; Pred. No. 51;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      4 KKICKMEKCSYFNVVNS 21
      1 : | : | : | : | : |
Db      21 KVLCDVNCACVYVNPFS 38

```

```

RESULT 15
US-09-708-427-4855
: Sequence 4855, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4855
: LENGTH: 256
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc-feature
: LOCATION: 1..256
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc-feature
: LOCATION: 1..256
: OTHER INFORMATION: Ceres Seq. ID 1813256
US-09-708-427-4855

```

```

Query Match          37.6%; Score 41; DB 5; Length 256;
Best Local Similarity 38.9%; Pred. No. 54;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      4 KKICKMEKCSYFNVVNS 21
      1 : | : | : | : | : |
Db      36 KVLCDVNCACVYVNPFS 53

```

Search completed: January 29, 2002, 10:58:10  
Job time: 2506 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:01 : Search time 1760.55 seconds  
(without alignments)  
3.627 Million cell updates/sec

Title: US-09-763-397A-6  
Perfect score: 127  
Sequence: 1 KPDELDYENDIEKICKMEKCS 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main: \*  
1: /cgn2\_6/ptodata/2/paa/PCRTUS.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082.COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083.COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084.COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085.COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086.COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087.COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088.COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089.COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090.COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091.COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092.COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093.COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094.COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095.COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096.COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097.COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US098.COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US099.COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US60.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	100.0	23	21	US-09-763-397A-6
2	127	100.0	24	3	US-07-610-222-33
3	127	100.0	24	8	US-08-480-512-55
4	127	100.0	350	21	US-09-763-397A-2
5	127	100.0	412	3	US-07-677-539B-6
6	127	100.0	412	3	US-07-727-636-6
7	127	100.0	412	13	US-08-948-885-18
8	127	100.0	423	13	US-08-933-929A-1
9	127	100.0	424	3	US-07-842-694-2

10	122	96.1	388	3	US-07-677-539B-5	Sequence 5, Appl1
11	122	96.1 <td>388</td> <td>3</td> <td>US-07-727-636-5</td> <td>Sequence 5, Appl1</td>	388	3	US-07-727-636-5	Sequence 5, Appl1
12	121	95.3	396	22	US-08-820-843A-31	Sequence 31, Appl1
13	121	95.3	424	11	US-08-760-797-1	Sequence 1, Appl1
14	121	95.3	424	11	US-08-760-797-3	Sequence 3, Appl1
15	121	95.3	424	13	US-08-903-084-1	Sequence 1, Appl1
16	121	95.3	424	13	US-08-903-084-3	Sequence 3, Appl1
17	121	95.3	424	13	US-08-932-929-1	Sequence 1, Appl1
18	121	95.3	424	13	US-08-932-929-3	Sequence 3, Appl1
19	121	95.3	424	13	US-08-932-929A-3	Sequence 3, Appl1
20	96	75.6	19	6	US-08-284-439A-42	Sequence 42, Appl1
21	96	75.6	19	13	US-08-971-790-42	Sequence 42, Appl1
22	75	59.1	20	3	US-07-947-033-19	Sequence 19, Appl1
23	75	59.1	20	5	US-08-119-694-19	Sequence 19, Appl1
24	75	59.1	20	5	US-08-119-694B-19	Sequence 19, Appl1
25	71	55.9	21	21	US-09-763-397A-7	Sequence 7, Appl1
26	62	48.8	14	3	US-07-947-033-24	Sequence 24, Appl1
27	62	48.8	14	5	US-08-119-694-24	Sequence 24, Appl1
28	62	48.8	14	5	US-08-119-694B-24	Sequence 24, Appl1
29	57	44.9	55	3	US-07-947-033-3	Sequence 3, Appl1
30	57	44.9	55	3	US-07-947-033-4	Sequence 4, Appl1
31	57	44.9	55	3	US-07-947-033-5	Sequence 5, Appl1
32	57	44.9	55	3	US-08-119-694-3	Sequence 3, Appl1
33	57	44.9	55	5	US-08-119-694-4	Sequence 4, Appl1
34	57	44.9	55	5	US-08-119-694-5	Sequence 5, Appl1
35	57	44.9	55	5	US-08-119-694B-3	Sequence 3, Appl1
36	57	44.9	55	5	US-08-119-694B-4	Sequence 4, Appl1
37	57	44.9	55	5	US-08-119-694B-5	Sequence 5, Appl1
38	53	41.7	21	1	PCT-US00-24802-3875	Sequence 3875, Ap
39	53	41.7	21	1	PCT-US00-33549-68	Sequence 68, Appl1
40	53	41.7	21	1	PCT-US00-34318-27	Sequence 27, Appl1
41	53	41.7	21	1	PCT-US00-35516-12	Sequence 12, Appl1
42	53	41.7	21	1	PCT-US94-04832A-48	Sequence 48, Appl1
43	53	41.7	21	1	PCT-US99-13923-41	Sequence 41, Appl1
44	53	41.7	21	1	PCT-US99-13959-78	Sequence 78, Appl1
45	53	41.7	21	5	US-08-121-101-17	Sequence 17, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-763-397A-6  
Sequence 6, Application US/09763397A  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America, as represented by the  
APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
APPLICANT: Control and Prevention  
APPLICANT: Lal, Altaf A.  
APPLICANT: Ping Shi, Ya  
APPLICANT: Hashain, Seyed E.  
TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
FILE REFERENCE: 6395-57049  
CURRENT APPLICATION NUMBER: US/09/763,397A  
CURRENT FILING DATE: 2001-02-16  
PRIOR APPLICATION NUMBER: US 60/097,703  
PRIOR FILING DATE: 1998-08-21  
PRIOR APPLICATION NUMBER: PCT / US99/18869  
PRIOR FILING DATE: 1999-08-19  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 6  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Plasmodium falciparum  
US-09-763-397A-6

Query Match 100.0% Score 127, DB 21: Length 23:  
Best Local Similarity 100.0% Pred. No. 2.5e-11:  
Matches 23: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
QY 1 KPDELDYENDIEKICKMEKCS 23

Db 1 KPKDELDYENDIEKKICKMEKCS 23

## RESULT 2

US-07-610-222-33

Sequence 33 Application US/07610222  
GENERAL INFORMATION:  
APPLICANT: Jung, Gunther  
APPLICANT: Rammensee, Hans-Georg  
APPLICANT: Schild, Hans-Jorg  
APPLICANT: Deres, Karl  
APPLICANT: Wiesmuller, Karl-Heinz  
TITLE OF INVENTION: Synthetic Vaccine for the Specific  
INDUCTION OF CYTOTOXIC T-LYMPHOCYTES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESS: Dunner  
STREET: 1300 I Street, N.W. Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/610,222  
FILING DATE: 19901108  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 39 37 412.2  
FILING DATE: 10-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavin Jr., Lawrence M.  
REGISTRATION NUMBER: 30,768  
REFERENCE/DOCKET NUMBER: 2481.1003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-610-222-33

Query Match 100.0%; Score 127; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2,6e-11;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KPKDELDYENDIEKKICKMEKCS 23  
Db 1 KPKDELDYENDIEKKICKMEKCS 23

## RESULT 3

US-08-480-512-55

Sequence 55 Application US/08480512  
GENERAL INFORMATION:  
APPLICANT: Jung, Gunther  
APPLICANT: Wiesmuller, Karl-Heinz  
APPLICANT: Metzger, Jorg  
APPLICANT: Buhning, Hans-Jorg  
APPLICANT: Becker, Gerhard  
APPLICANT: Bessler, Wolfgang  
APPLICANT: Hess, Gunther

APPLICANT: Rammensee, Hans-Georg  
APPLICANT: Schild, Hans-Jorg  
APPLICANT: Deres, Karl  
TITLE OF INVENTION: Membrane Anchor/Active Compound  
INDUCTION OF CYTOTOXIC T-LYMPHOCYTES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESS: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,512  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/387,624  
FILING DATE:  
APPLICATION NUMBER: US/08/084,091A  
FILING DATE: 30-JUN-1993  
APPLICATION NUMBER: US/07/588,794  
FILING DATE: 27-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/427,914  
FILING DATE: 24-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/229,770  
FILING DATE: 01-AUG-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/06/876,479  
FILING DATE: 20-JUN-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/340,833  
FILING DATE: 20-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/966,603  
FILING DATE: 26-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/610,222  
FILING DATE: 08-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P3522512.2  
FILING DATE: 24-JUN-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P3546150.0  
FILING DATE: 27-DEC-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P3813821.2  
FILING DATE: 22-APR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P3937412.2  
FILING DATE: 10-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Flesher, Raz E.  
REGISTRATION NUMBER: 34,331  
REFERENCE/DOCKET NUMBER: 02481.0312-04000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-480-512-55

Query Match 100.0%; Score 127; DB 8; Length 24;

Best Local Similarity 100.0%; Pred. No. 2.6e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPDELVDYNDIEKKICKMEKCS 23

DB 1 KPDELVDYNDIEKKICKMEKCS 23

RESULT 4

US-09-763-397A-2

Sequence 2, Application US/09763397A

GENERAL INFORMATION:

APPLICANT: The Government of the United States of America, as represented by the

APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis

APPLICANT: Control and Prevention

APPLICANT: Lal, Altaf A.

APPLICANT: Hashain, Seyed E.

TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci

FILE REFERENCE: 6395-57049

CURRENT APPLICATION NUMBER: US/09/763,397A

CURRENT FILING DATE: 2001-02-16

PRIOR APPLICATION NUMBER: US 60/097,703

PRIOR FILING DATE: 1998-08-21

PRIOR APPLICATION NUMBER: PCT / US99/18869

PRIOR FILING DATE: 1999-08-19

NUMBER OF SEQ ID NOS: 26

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 350

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Recombinant DNA/Protein

US-09-763-397A-2

Query Match 100.0%; Score 127; DB 21; Length 350;

Best Local Similarity 100.0%; Pred. No. 6.9e-10;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPDELVDYNDIEKKICKMEKCS 23

DB 51 KPDELVDYNDIEKKICKMEKCS 73

RESULT 5

US-07-677-539B-6

Sequence 6, Application US/07677539B

GENERAL INFORMATION:

APPLICANT: LAL, ALTA F.

APPLICANT: GOLDMAN, IRA F.

TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM

TITLE OF INVENTION: REICHENOMI AND VACCINE FOR HUMAN MALARIA

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN, DARBY &amp; CUSHMAN

STREET: 1615 L STREET, N.W.

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

FILING DATE: 19911205

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: SCOTT, WATSON T.

REGISTRATION NUMBER: 26,581

REFERENCE/DOCKET NUMBER: 5683/91540

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-861-3067

TELEFAX: 202-822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 412 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-677-539B-6

Query Match 100.0%; Score 127; DB 3; Length 412;

Best Local Similarity 100.0%; Pred. No. 8.5e-10;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPDELVDYNDIEKKICKMEKCS 23

DB 368 KPDELVDYNDIEKKICKMEKCS 390

RESULT 6

US-07-727-636-6

Sequence 6, Application US/07727636

GENERAL INFORMATION:

APPLICANT: LAL, ALTA F.

APPLICANT: GOLDMAN, IRA F.

TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM

TITLE OF INVENTION: REICHENOMI AND VACCINE FOR HUMAN MALARIA

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN, DARBY &amp; CUSHMAN

STREET: 1615 L STREET, N.W.

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/727,636

FILING DATE: 19910710

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: SCOTT, WATSON T.

REGISTRATION NUMBER: 26,581

REFERENCE/DOCKET NUMBER: 5683/91540

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-861-3067

TELEFAX: 202-822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 412 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-727-636-6

Query Match 100.0%; Score 127; DB 3; Length 412;

Best Local Similarity 100.0%; Pred. No. 8.5e-10;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPDELVDYNDIEKKICKMEKCS 23

DB 368 KPDELVDYNDIEKKICKMEKCS 390

RESULT 6

US-07-727-636-6

Sequence 6, Application US/07727636

GENERAL INFORMATION:

APPLICANT: LAL, ALTA F.

APPLICANT: GOLDMAN, IRA F.

TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM

TITLE OF INVENTION: REICHENOMI AND VACCINE FOR HUMAN MALARIA

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN, DARBY &amp; CUSHMAN

STREET: 1615 L STREET, N.W.

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/727,636

FILING DATE: 19910710

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: SCOTT, WATSON T.

REGISTRATION NUMBER: 26,581

REFERENCE/DOCKET NUMBER: 5683/91540

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-861-3067

TELEFAX: 202-822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 412 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-727-636-6

Query Match 100.0%; Score 127; DB 3; Length 412;

Best Local Similarity 100.0%; Pred. No. 8.5e-10;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPDELVDYNDIEKKICKMEKCS 23

DB 368 KPDELVDYNDIEKKICKMEKCS 390

RESULT 6

US-07-727-636-6

Sequence 6, Application US/07727636

GENERAL INFORMATION:

APPLICANT: LAL, ALTA F.

APPLICANT: GOLDMAN, IRA F.

TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM

TITLE OF INVENTION: REICHENOMI AND VACCINE FOR HUMAN MALARIA

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN, DARBY &amp; CUSHMAN

STREET: 1615 L STREET, N.W.

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/727,636

Best Local Similarity 100.0%; Pred. No. 8.5e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KPKDELVDYNDIEKKICKMEKCS 23

Db 368 KPKDELVDYNDIEKKICKMEKCS 390

RESULT 7

US-08-948-885-18

Sequence 18, Application US/08948885  
GENERAL INFORMATION:  
APPLICANT: Jessell, Thomas M. and Avihu Klar  
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A  
TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/948,885  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/313,288  
FILING DATE: January 5, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0526  
TELEX:  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 412 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-948-885-18

Query Match 100.0%; Score 127; DB 13; Length 412;  
Best Local Similarity 100.0%; Pred. No. 8.5e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KPKDELVDYNDIEKKICKMEKCS 23

Db 368 KPKDELVDYNDIEKKICKMEKCS 390

RESULT 8

US-08-932-929A-1

Sequence 1, Application US/08932929A  
GENERAL INFORMATION:  
APPLICANT: De Wilde, Michel  
APPLICANT: Cohen, Joseph  
TITLE OF INVENTION: Hybrid Protein Between CS  
TITLE OF INVENTION: from Plasmodium and HbsAg  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/932,929A

FILING DATE: 18-SEPT-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/663,371

FILING DATE: 13-JUNE-1996

ATTORNEY/AGENT INFORMATION:

NAME: Baumeister, Kirk

REGISTRATION NUMBER: 33,833

REFERENCE/DOCKET NUMBER: B45015-1FWC2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5096

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 423 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-932-929A-1

Query Match 100.0%; Score 127; DB 13; Length 423;  
Best Local Similarity 100.0%; Pred. No. 8.7e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KPKDELVDYNDIEKKICKMEKCS 23

Db 162 KPKDELVDYNDIEKKICKMEKCS 184

RESULT 9

US-07-842-694-2

Sequence 2, Application US/07842694  
GENERAL INFORMATION:  
APPLICANT: Cohen, Joseph  
APPLICANT: De Wilde, Michel  
TITLE OF INVENTION: Malaria Vaccine Antigen  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: P.O. Box 1539 / Corporate Patents  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/842,694  
FILING DATE: 19920227  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Canter, Carol G.  
REGISTRATION NUMBER: 31151  
REFERENCE/DOCKET NUMBER: B45015  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 424 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-07-842-694-2

Query Match 100.0%; Score 127; DB 3; Length 424;  
Best Local Similarity 100.0%; Pred. No. 8.8e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPKDELVDYENDIEKKICKMEKCS 23  
DB 163 KPKDELVDYENDIEKKICKMEKCS 185

## RESULT 10

US-07-677-539B-5  
Sequence 5, Application US/07677539B  
GENERAL INFORMATION:  
APPLICANT: LAL, ALTAI A.  
APPLICANT: GOLDMAN, IRA F.  
TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM  
TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1615 L STREET, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/677,539B  
FILING DATE: 19911205  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: SCOTT, WATSON T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: 5683/91540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3067  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-677-539B-5

Query Match 96.1%; Score 122; DB 3; Length 388;  
Best Local Similarity 91.3%; Pred. No. 4.3e-09;  
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPKDELVDYENDIEKKICKMEKCS 23  
DB 344 KPKDELVDYENDIEKKICKMEKCS 366

RESULT 11  
US-07-727-636-5

Sequence 5, Application US/07727636  
GENERAL INFORMATION:  
APPLICANT: LAL, ALTAI A.  
APPLICANT: GOLDMAN, IRA F.  
TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM  
TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1615 L STREET, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/727,636  
FILING DATE: 19910710  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: SCOTT, WATSON T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: 5683/91540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3067  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-727-636-5

Query Match 96.1%; Score 122; DB 3; Length 388;  
Best Local Similarity 91.3%; Pred. No. 4.3e-09;  
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPKDELVDYENDIEKKICKMEKCS 23  
DB 344 KPKDELVDYENDIEKKICKMEKCS 366

RESULT 12  
US-09-820-843A-31  
Sequence 31, Application US/09820843A  
GENERAL INFORMATION:  
APPLICANT: Council of Scientific and Industrial Research  
TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PRO  
TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES  
FILE REFERENCE: 063915  
CURRENT APPLICATION NUMBER: US/09/820,843A  
CURRENT FILING DATE: 2001-03-30  
NUMBER OF SEQ ID NOS: 118  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 31  
LENGTH: 396  
TYPE: PRT  
ORGANISM: Plasmodium falciparum  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Circumsporozone (CS) protein  
OTHER INFORMATION: g114493889  
US-09-820-843A-31

Query Match 95.3%; Score 121; DB 22; Length 396;  
Best Local Similarity 95.7%; Pred. No. 6.1e-09;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPDELVDYNDIEKKICKMEKCS 23  
||||| |||||||  
DB 352 KPDELVDYNDIEKKICKMEKCS 374

## RESULT 13

US-08-760-797-1  
; Sequence 1, Application US/08760797  
; GENERAL INFORMATION:  
; APPLICANT: De Wilde, Michel  
; APPLICANT: Cohen, Joseph  
; TITLE OF INVENTION: HYBRID PROTEIN BETWEEN CS FROM  
; TITLE OF INVENTION: PLASMODIUM AND HBSAG  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: SmithKline Beecham Corp./Corporate  
; STREET: P.O. Box 1539 - UW2220  
; CITY: King of Prussia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/760,797  
; FILING DATE: 04-DEC-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/442,612  
; FILING DATE: 17-MAY-1995  
; APPLICATION NUMBER: US 08/244,085  
; FILING DATE: 21-OCT-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: B45015C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610 270-5096  
; TELEFAX: 610 270-5090  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 424 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-760-797-1

Query Match 95.3%; Score 121; DB 11; Length 424;  
Best Local Similarity 95.7%; Pred. No. 6.6e-09;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPDELVDYNDIEKKICKMEKCS 23  
||||| |||||||  
DB 159 KPDELVDYNDIEKKICKMEKCS 181

## RESULT 14

US-08-760-797-3  
; Sequence 3, Application US/08760797  
; GENERAL INFORMATION:  
; APPLICANT: De Wilde, Michel

APPLICANT: Cohen, Joseph  
; TITLE OF INVENTION: HYBRID PROTEIN BETWEEN CS FROM  
; TITLE OF INVENTION: PLASMODIUM AND HBSAG  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: SmithKline Beecham Corp./Corporate  
; ADDRESSEE: Intellectual Property  
; STREET: P.O. Box 1539 - UW2220  
; CITY: King of Prussia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19406-0939

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/760,797  
; FILING DATE: 04-DEC-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/442,612  
; FILING DATE: 17-MAY-1995  
; APPLICATION NUMBER: US 08/244,085  
; FILING DATE: 21-OCT-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: B45015C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610 270-5096  
; TELEFAX: 610 270-5090  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 424 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-760-797-3

Query Match 95.3%; Score 121; DB 11; Length 424;  
Best Local Similarity 95.7%; Pred. No. 6.6e-09;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KPDELVDYNDIEKKICKMEKCS 23  
||||| |||||||  
DB 159 KPDELVDYNDIEKKICKMEKCS 181

## RESULT 15

US-08-903-084-1  
; Sequence 1, Application US/08903084  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Francotte, Myriam  
; APPLICANT: Kummer, Suzanne  
; APPLICANT: Slaoui, Moncef  
; APPLICANT: Wijndale, Frans  
; TITLE OF INVENTION: VACCINE FORMULATIONS CONTAINING  
; TITLE OF INVENTION: 3-O-DEACETYLATED MONOPHOSPHORYL LIPID A AND A CARRIER  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: U.S.A.  
; ZIP: 19046  
; COMPUTER READABLE FORM:



```

1      MEDIUM TYPE: Diskette
2      COMPUTER: IBM Compatible
3      OPERATING SYSTEM: DOS
4      SOFTWARE: FASTSEQ for Windows Version 2.0
5      CURRENT APPLICATION DATA:
6      APPLICATION NUMBER: US/08/903,084
7      FILING DATE: 17-JUL-1997
8      CLASSIFICATION: 424
9      PRIOR APPLICATION DATA:
10     APPLICATION NUMBER: 08/741,575
11     FILING DATE: 30-OCT-1996
12     APPLICATION NUMBER: 08/303,542
13     FILING DATE: 09-SEP-1994
14     ATTORNEY/AGENT INFORMATION:
15     NAME: Kerekes, Zoltan
16     REGISTRATION NUMBER: 98,938
17     REFERENCE/DOCKET NUMBER: B45101
18     TELECOMMUNICATION INFORMATION:
19     TELEPHONE: 610-270-5024
20     TELEFAX: 610-270-5090
21     TELEX:
22     INFORMATION FOR SEQ ID NO: 1:
23     SEQUENCE CHARACTERISTICS:
24     LENGTH: 424 amino acids
25     TYPE: amino acid
26     STRANDEDNESS: single
27     TOPOLOGY: linear
28     MOLECULE TYPE: protein
29     US-08-903-084-1

```

Query Match	95.3%	Score 121	DB 13	Length 424
Best Local Similarity	95.7%	Pred. No. 6	6e-09	
Matches 22	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1	KPKDELVDYNDIEKKIKCKMEKCS	23	
Db	159	KPKDELVDYNDIEKKIKCKMEKCS	181	

Search completed: January 29, 2002, 10:56:01  
Job time: 2397 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:09 ; Search time 120.95 seconds  
(without alignments)  
6.959 Million cell updates/sec

Title: US-09-763-397a-6

Perfect score: 127  
Sequence: 1 KPKDELNDYENDIEKICKMEKCS 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCR\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	36.2	867	5	US-09-708-427-5053
2	46	36.2	872	5	US-09-708-427-5052
3	46	36.2	920	5	US-09-708-427-5051
4	45.5	35.8	1124	5	US-09-897-516-6570
5	45	35.4	922	6	US-10-018-105-4
6	44	34.6	416	5	US-09-708-427-7358
7	44	34.6	533	5	US-09-708-427-7357
8	44	34.6	550	5	US-09-708-427-19153
9	44	34.6	553	5	US-09-708-427-7356
10	44	34.6	558	5	US-09-708-427-19152
11	44	34.6	577	5	US-09-708-427-19151
12	44	34.6	607	5	US-09-815-242-13379
13	44	34.6	607	5	US-09-815-242-13682
14	44	34.6	772	5	US-09-708-427-14304
15	44	34.6	803	5	US-09-708-427-14303
16	44	34.6	804	5	US-09-708-427-24127
17	44	34.6	807	5	US-09-708-427-24126
18	44	34.6	819	5	US-09-708-427-20310
19	44	34.6	821	5	US-09-708-427-24125
20	44	34.6	837	5	US-09-708-427-14302
21	43.5	34.3	190	5	US-09-708-427-35295
22	43.5	34.3	297	5	US-09-708-427-35294
23	43.5	34.3	361	5	US-09-708-427-35293
24	43	33.9	133	5	US-09-708-427-33267
25	43	33.9	181	5	US-09-708-427-33266
26	43	33.9	207	5	US-09-708-427-33265

27	43	33.9	272	5	US-09-708-427-56477	Sequence 56477, A
28	43	33.9	309	5	US-09-708-427-56476	Sequence 56476, A
29	43	33.9	342	5	US-09-708-427-56475	Sequence 56475, A
30	42	33.1	32	7	US-60-323-991-30	Sequence 30, Appl
31	42	33.1	150	5	US-09-708-427-39156	Sequence 39156, A
32	42	33.1	163	5	US-09-620-3948-5279	Sequence 5279, Ap
33	42	33.1	254	5	US-09-815-242-11461	Sequence 11461, A
34	42	33.1	467	5	US-09-611-526-3365	Sequence 3265, A
35	41.5	32.7	1204	5	US-09-708-427-11575	Sequence 11575, A
36	41.5	32.7	1245	5	US-09-708-427-11574	Sequence 11574, A
37	41.5	32.7	1249	5	US-09-708-427-11573	Sequence 11573, A
38	41	32.3	232	5	US-09-815-242-5560	Sequence 5560, Ap
39	41	32.3	232	5	US-09-815-242-11666	Sequence 11666, A
40	41	32.3	232	5	US-09-815-242-12569	Sequence 12569, A
41	41	32.3	270	5	US-09-620-3948-1049	Sequence 1049, Ap
42	41	32.3	272	5	US-09-620-3948-1048	Sequence 1048, Ap
43	41	32.3	336	5	US-09-620-3948-1047	Sequence 1047, Ap
44	41	32.3	402	5	US-09-708-427-5893	Sequence 5893, Ap
45	41	32.3	420	5	US-09-708-427-5892	Sequence 5892, Ap

## ALIGNMENTS

```
RESULT 1
US-09-708-427-5053
; Sequence 5053, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO. 5053
; LENGTH: 867
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..867
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..867
; OTHER INFORMATION: Ceres Seq. ID 1813621
US-09-708-427-5053

Query Match 36.2%; Score 46; DB 5; Length 867;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

3 KPELDYENDIEKICKME 20
Db 816 RDTDSYINDIEKICLSOE 833

RESULT 2
US-09-708-427-5052
; Sequence 5052, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO. 5052
; LENGTH: 872
```

```
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..872
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..872
; OTHER INFORMATION: Ceres Seq. ID 1813620
US-09-708-427-5052

Query Match          36.2%; Score 46; DB 5; Length 872;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 KDELVDYNDIEKKICKME 20
   : | | | | | | | | | |
Db 821 RDTDSYINDIEKICLSOE 838

RESULT 3
US-09-708-427-5051
; Sequence 5051, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5051
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..920
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..920
; OTHER INFORMATION: Ceres Seq. ID 1813619
US-09-708-427-5051

Query Match          36.2%; Score 46; DB 5; Length 920;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 KDELVDYNDIEKKICKME 20
   : | | | | | | | | | |
Db 869 RDTDSYINDIEKICLSOE 886

RESULT 4
US-09-897-516-6570
; Sequence 6570, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; OTHER INFORMATION: Xaa is any amino acid
; PRIOR APPLICATION NUMBER: US 60/215, 161
```

```
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 6570
; LENGTH: 1124
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516-6570

Query Match          35.8%; Score 45.5; DB 5; Length 1124;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 5 ELVDYEN--DIEKKICKMEK 21
   : | | | | : | | | | |
Db 208 EMDAENIAAEIRKKLCYQEK 227

RESULT 5
US-10-018-105-4
; Sequence 4, Application US/10018105
; GENERAL INFORMATION:
; APPLICANT: ROEMER, Terry
; APPLICANT: BUSSEY, Howard
; APPLICANT: DAVISON, John
; TITLE OF INVENTION: IDENTIFICATION OF CANDIDA ALBICANS ESSENTIAL FUNGAL
; TITLE OF INVENTION: SPECIFIC GENES AND USE THEREOF IN ANTIFUNGAL DRUG
; FILE REFERENCE: 10182-015
; CURRENT APPLICATION NUMBER: US/10/018,105
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: PCT/CA00/00533
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 60/132,878
; PRIOR FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 922
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-018-105-4

Query Match          35.4%; Score 45; DB 6; Length 922;
Best Local Similarity 36.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 3 KDELVDYNDIEKKICKMEK 21
   : | | | : | | | : | | |
Db 41 EDELKSELESEVVKSEK 59

RESULT 6
US-09-708-427-7358
; Sequence 7358, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7358
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..416
; OTHER INFORMATION: Xaa is any amino acid
```



Query Match 34.6%; Score 44; DB 5; Length 558;  
Best Local Similarity 47.6%; Pred. No. 1.5e+02;  
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPKDELVDNDIEKKICKMEK 21  
||| ||| : | : ||  
Db 267 KPKKELDDQDKAKAVNGEEK 267

RESULT 11  
US-09-708-427-19151

; Sequence 19151, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19151  
; LENGTH: 577  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..577  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc.feature  
; LOCATION: 1..577  
; OTHER INFORMATION: Ceres Seq. ID 1835492  
US-09-708-427-19151

Query Match 34.6%; Score 44; DB 5; Length 577;  
Best Local Similarity 47.6%; Pred. No. 1.5e+02;  
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPKDELVDNDIEKKICKMEK 21  
||| ||| : | : ||  
Db 286 KPKKELDDQDKAKAVNGEEK 306

RESULT 12  
US-09-815-242-13379

; Sequence 13379, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 13379  
; LENGTH: 607  
; TYPE: PRP  
; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13379

Query Match 34.6%; Score 44; DB 5; Length 607;  
Best Local Similarity 38.1%; Pred. No. 1.6e+02;  
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 KPKDELVDNDIEKKICKMEK 21  
||| ||| : | : ||  
Db 498 KRKEVDLRNEVDQALFATEK 518

RESULT 13  
US-09-815-242-13682

; Sequence 13682, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 13682  
; LENGTH: 607  
; TYPE: PRP  
; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13682

Query Match 34.6%; Score 44; DB 5; Length 607;  
Best Local Similarity 38.1%; Pred. No. 1.6e+02;  
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 KPKDELVDNDIEKKICKMEK 21  
||| ||| : | : ||  
Db 498 KRKEVDLRNEVDQALFATEK 518

RESULT 14  
US-09-708-427-14304

; Sequence 14304, Application US/09708427  
; GENERAL INFORMATION:

```

; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243p
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14304
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..772
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..772
; OTHER INFORMATION: Ceres Seq. ID 1827597
US-09-708-427-14304

```

```

Query Match          34.6%: Score 44; DB 5; Length 772;
Best Local Similarity 44.4%: Pred. No. 2e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

```

Oy 4 DELDYENDIEKICKMEK 21
   1: | : | 1: | : |
Db 616 DDPDLQNTFFYKLCYK 633

```

```

RESULT 15
US-09-708-427-14303
; Sequence 14303, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243p
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14303
; LENGTH: 803
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..803
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..803
; OTHER INFORMATION: Ceres Seq. ID 1827596
US-09-708-427-14303

```

```

Query Match          34.6%: Score 44; DB 5; Length 803;
Best Local Similarity 44.4%: Pred. No. 2.1e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

```

Oy 4 DELDYENDIEKICKMEK 21
   1: | : | 1: | : |
Db 647 DDPDLQNTFFYKLCYK 664

```

Search completed: January 29, 2002, 10:58:09  
 Job time: 2505 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:56:00 : Search time 1760.55 seconds  
(without alignments)  
2.050 Million cell updates/sec

Title: US-09-763-397A-5  
Perfect score: 75  
Sequence: 1 KHKKLKQPGDGNP 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/US085.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US086.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US087.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US088.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US089.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US090.COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US091.COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US092.COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US093.COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US094.COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US095.COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US096.COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US097.COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US098.COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US099.COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US100.COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US101.COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US102.COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US103.COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US104.COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US105.COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US106.COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US107.COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US108.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	13	US-09-763-397A-5	Sequence 5, Appli
2	75	100.0	350	US-09-763-397A-2	Sequence 2, Appli
3	75	100.0	412	US-07-677-539B-6	Sequence 6, Appli
4	75	100.0	412	US-07-727-636-6	Sequence 6, Appli
5	75	100.0	412	US-08-948-885-18	Sequence 18, Appli
6	69	92.0	20	US-07-947-033-21	Sequence 21, Appli
7	69	92.0	20	US-08-119-694-21	Sequence 21, Appli
8	69	92.0	20	US-08-119-694B-21	Sequence 21, Appli
9	69	92.0	396	US-09-820-843A-31	Sequence 31, Appli

10	51	68.0	388	3	US-07-677-539B-5	Sequence 5, Appli
11	51	68.0	388	3	US-07-727-636-5	Sequence 5, Appli
12	47	62.7	102	1	PCT-US01-14826-223	Sequence 223, App
13	47	62.7	102	1	PCT-US01-23861-2	Sequence 2, Appli
14	47	62.7	102	1	PCT-US99-09970-2	Sequence 2, Appli
15	47	62.7	102	17	US-09-306-111-2	Sequence 2, Appli
16	47	62.7	102	22	US-09-834-366-22614	Sequence 2, Appli
17	47	62.7	102	22	US-09-834-366-24566	Sequence 24566, A
18	47	62.7	102	24	US-60-197-873-22614	Sequence 22614, A
19	47	62.7	102	24	US-60-197-873-24566	Sequence 24566, A
20	47	62.7	102	24	US-60-222-039-2	Sequence 2, Appli
21	47	62.7	118	1	PCT-US01-14826-655	Sequence 655, App
22	47	62.7	126	21	US-09-757-026-455	Sequence 455, App
23	46	61.3	101	11	US-08-700-408-68	Sequence 68, Appli
24	46	61.3	101	18	US-09-485-421-2	Sequence 2, Appli
25	46	61.3	101	18	US-09-485-421A-2	Sequence 2, Appli
26	43	57.3	108	24	US-60-207-583-602	Sequence 602, App
27	43	57.3	452	16	US-09-248-796-16171	Sequence 16171, A
28	41	54.7	129	19	US-09-540-236-2919	Sequence 2919, Ap
29	41	54.7	157	24	US-60-178-307-2326	Sequence 2326, Ap
30	41	54.7	157	24	US-60-196-713-2331	Sequence 2331, Ap
31	41	54.7	208	24	US-60-128-476-3431	Sequence 3431, Ap
32	40.5	54.0	1332	1	PCT-US01-08631-39492	Sequence 39492, A
33	40	53.3	75	22	US-09-834-366-21504	Sequence 21504, A
34	40	53.3	75	24	US-60-197-873-21504	Sequence 21504, A
35	40	53.3	86	21	US-09-758-474-1239	Sequence 1239, Ap
36	40	53.3	89	13	US-08-905-135-478	Sequence 478, App
37	40	53.3	96	21	US-09-732-210-707	Sequence 707, App
38	40	53.3	96	24	US-60-169-340-707	Sequence 707, App
39	40	53.3	101	22	US-09-834-366-17600	Sequence 17600, A
40	40	53.3	101	24	US-60-197-873-17600	Sequence 17600, A
41	40	53.3	109	24	US-60-215-161-4572	Sequence 4572, Ap
42	40	53.3	164	20	US-09-628-359-38	Sequence 38, Appli
43	40	53.3	209	20	US-60-324-109-19086	Sequence 19086, A
44	40	53.3	211	20	US-09-628-359-37	Sequence 37, Appli
45	40	53.3	297	24	US-60-324-109-18985	Sequence 18985, A

ALIGNMENTS

RESULT 1  
US-09-763-397A-5  
Sequence 5, Application US/09763397A  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America, as represented by the  
APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
APPLICANT: Control and Prevention  
APPLICANT: Lal, Altaf A.  
APPLICANT: Ping Shl, Ya  
APPLICANT: Hashain, Seyed E.  
TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
FILE REFERENCE: 6395-57049  
CURRENT APPLICATION NUMBER: US/09/763,397A  
CURRENT FILING DATE: 2001-02-16  
PRIOR FILING DATE: 1998-08-21  
PRIOR APPLICATION NUMBER: PCT / US99/18869  
PRIOR FILING DATE: 1999-08-19  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Plasmodium falciparum  
US-09-763-397A-5

Query Match 100.0% Score 75; DB 21; Length 13;  
Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KHKKLKQPGDGNP 13

Db 1 KHKRLKOPGDSNP 13

RESULT 2  
US-09-763-397A-2

; Sequence 2, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for Dis  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seyed E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Falci  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763,397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097,703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 2  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Recombinant DNA/Protein  
US-09-763-397A-2

Query Match 100.0%; Score 75; DB 21; Length 350;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHKRLKOPGDSNP 13  
Db 29 KHKRLKOPGDSNP 41

RESULT 3  
US-07-677-539B-6

; Sequence 6, Application US/07677539B  
; GENERAL INFORMATION:  
; APPLICANT: LAL, ALTAf A.  
; APPLICANT: GOLDMAN, IRA F.  
; TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM  
; TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L STREET, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/677,539B  
; FILING DATE: 19911205  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SCOTT, WATSON T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: 5683/91540  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3067

; TELEFAX: 202-822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 412 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-07-677-539B-6

Query Match 100.0%; Score 75; DB 3; Length 412;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHKRLKOPGDSNP 13  
Db 109 KHKRLKOPGDSNP 121

RESULT 4  
US-07-727-636-6

; Sequence 6, Application US/07727636  
; GENERAL INFORMATION:  
; APPLICANT: LAL, ALTAf A.  
; APPLICANT: GOLDMAN, IRA F.  
; TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM  
; TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L STREET, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/727,636  
; FILING DATE: 19910710  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SCOTT, WATSON T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: 5683/91540  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3067  
; TELEFAX: 202-822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 412 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-07-727-636-6

Query Match 100.0%; Score 75; DB 3; Length 412;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHKRLKOPGDSNP 13  
Db 109 KHKRLKOPGDSNP 121

RESULT 5  
US-08-948-885-18  
; Sequence 18, Application US/08948885  
; GENERAL INFORMATION:  
; APPLICANT: Jessell, Thomas M. and Avihu Klar  
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A  
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/948,885  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/313,288  
; FILING DATE: January 5, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0526  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 412 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-948-885-18  
  
Query Match 100.0%; Score 75; DB 13; Length 412;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 KKKLKOPGDGPNP 13  
Db 109 KKKLKOPGDGPNP 121  
  
RESULT 6  
US-07-947-033-21  
; Sequence 21, Application US/07947033  
; GENERAL INFORMATION:  
; APPLICANT: Cerami, Carla  
; APPLICANT: Frevert, Ute  
; APPLICANT: Siminis, Photini  
; APPLICANT: Nussenzweig, Victor  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING  
; TITLE OF INVENTION: HEPATOCYTE INVASION BY MALARIAL SPOOROZOITES  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Darby & Darby, P.C.  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10022  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage  
; COMPUTER: AST Premium II 386/33  
; OPERATING SYSTEM: DOS 3.3  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/947,033  
; FILING DATE: 19920917  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adda C. Gogoris, Esq.  
; REGISTRATION NUMBER: 29,714  
; REFERENCE/DOCKET NUMBER: 5986/07607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 527-7700  
; TELEFAX: (212) 753-6237  
; TELEX: 236687  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
; US-07-947-033-21  
  
Query Match 92.0%; Score 69; DB 3; Length 20;  
Best Local Similarity 92.3%; Pred. No. 0.00071;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
OY 1 KKKLKOPGDGPNP 13  
Db 3 KKKLKOPADGPNP 15  
  
RESULT 7  
US-08-119-694-21  
; Sequence 21, Application US/08119694  
; GENERAL INFORMATION:  
; APPLICANT: Cerami, Carla  
; APPLICANT: Frevert, Ute  
; APPLICANT: Siminis, Photini  
; APPLICANT: Nussenzweig, Victor  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING  
; TITLE OF INVENTION: HEPATOCYTE INVASION BY MALARIAL  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Darby & Darby, P.C.  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5, inch, 1.2 MB storage  
; COMPUTER: AST Premium II 386/33  
; OPERATING SYSTEM: DOS 3.3  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/119,694  
; FILING DATE: 10-Sept-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adda C. Gogoris, Esq.  
; REGISTRATION NUMBER: 29,714  
; REFERENCE/DOCKET NUMBER: 5986/17607-US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 527-7700  
; TELEFAX: (212) 753-6237  
; TELEX: 236687  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown

US-08-119-694-21

Query Match 92.0%; Score 69; DB 5; Length 20;  
Best Local Similarity 92.3%; Pred. No. 0.00071;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKLKQPGDGNP 13  
Db 3 KKKLKQPGDGNP 15

RESULT 8

US-08-119-694B-21

; Sequence 21, Application US/08119694B  
; GENERAL INFORMATION:  
; APPLICANT: Cerami, Carla  
; APPLICANT: Frevert, Ute  
; APPLICANT: Simis, Photini  
; APPLICANT: Nussenzweig, Victor  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: INHIBITING HEPATOCYTE INVASION  
; TITLE OF INVENTION: BY MALARIAL SPOROZOITES  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Darby & Darby, P.C.  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5, inch, 360 KB storage  
; OPERATING SYSTEM: DOS 3.3  
; SOFTWARE: Word Perfect 5.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/119,694B  
; FILING DATE: 10-Sept-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adda C. Gogoris, Esq.  
; REGISTRATION NUMBER: 29,714  
; REFERENCE/DOCKET NUMBER: 5986/17607-US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 527-7700  
; TELEFAX: (212) 753-6237  
; TELEX: 236687  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; US-08-119-694B-21

Query Match 92.0%; Score 69; DB 5; Length 20;  
Best Local Similarity 92.3%; Pred. No. 0.00071;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKLKQPGDGNP 13  
Db 3 KKKLKQPGDGNP 15

RESULT 9  
US-09-820-843A-31  
; Sequence 31, Application US/09820843A  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN  
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES  
; FILE REFERENCE: Q63915  
; CURRENT APPLICATION NUMBER: US/09/820,843A

; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 31  
; LENGTH: 396  
; TYPE: PR  
; ORGANISM: Plasmodium falciparum  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Circumsporozoite (CS) protein  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: g14493889  
US-09-820-843A-31

Query Match 92.0%; Score 69; DB 22; Length 396;  
Best Local Similarity 92.3%; Pred. No. 0.023;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKLKQPGDGNP 13  
Db 89 KKKLKQPGDGNP 101

RESULT 10

US-07-677-539B-5

; Sequence 5, Application US/07677539B  
; GENERAL INFORMATION:  
; APPLICANT: LAL, ALTA F.  
; APPLICANT: GOLDMAN, IRA F.  
; TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM  
; TITLE OF INVENTION: REICHENOWI AND VACCINE FOR HUMAN MALARIA  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1615 L STREET, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/677,539B  
; FILING DATE: 19911205  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SCOTT, WATSON T.  
; REGISTRATION NUMBER: 26,581  
; REFERENCE/DOCKET NUMBER: 5683/91540  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3067  
; TELEFAX: 202-822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 388 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-07-677-539B-5

Query Match 68.0%; Score 51; DB 3; Length 388;  
Best Local Similarity 75.0%; Pred. No. 17;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKLKQPGDGN 12  
|| ||||| |

Db 109 KHKLKOPGNDN 120

RESULT 11

US-07-727-636-5

Sequence 5, Application US/07727636

GENERAL INFORMATION:

APPLICANT: LAL, ALTAI A.

APPLICANT: GOLDMAN, IRA F.

TITLE OF INVENTION: CIRCUMSPOROZOITE PROTEIN OF PLASMODIUM

TITLE OF INVENTION: CIRCUMSPOROZOITE AND VACCINE FOR HUMAN MALARIA

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN, DARBY & CUSHMAN

STREET: 1615 L STREET, N.W.

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/727.636

FILING DATE: 19910710

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: SCOTT, WATSON T.

REGISTRATION NUMBER: 26,581

REFERENCE/DOCKET NUMBER: 5683/91540

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-861-3067

TELEFAX: 202-822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 388 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-727-636-5

Query Match 68.0%; Score 51; DB 3; Length 388;  
Best Local Similarity 75.0%; Pred. No. 17;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KHKLKOPGNDN 12

Db 109 KHKLKOPGNDN 120

RESULT 12

PCT-US01-14826-223

Sequence 223, Application PC/TUS0114826

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc

TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

FILE REFERENCE: 21272-103

CURRENT APPLICATION NUMBER: PCT/US01/14826

PRIOR FILING DATE: 2001-05-16

PRIOR APPLICATION NUMBER: 09/577,408

PRIOR APPLICATION NUMBER: 09/677,298<151> 2000-09-22

PRIOR APPLICATION NUMBER: 09/695,781<151> 2000-10-24

PRIOR APPLICATION NUMBER: 09/715,868<151> 2000-11-17

PRIOR APPLICATION NUMBER: 09/775,330<151> 2001-02-01

NUMBER OF SEQ ID NOS: 864

SOFTWARE: Custom

SEQ ID NO 223

TYPE: PRT

LENGTH: 102  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-14826-223

Query Match 62.7%; Score 47; DB 1; Length 102;  
Best Local Similarity 61.5%; Pred. No. 16;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KHKLKOPGNDN 13

Db 90 KHKLKOPGNDN 102

RESULT 13

PCT-US01-23861-2

Sequence 2, Application PC/TUS0123861

GENERAL INFORMATION:

APPLICANT: Mungert, William E.

APPLICANT: Kulkarni, Prakash

APPLICANT: Getzenberg, Robert R.

TITLE OF INVENTION: Expression of a GAGE/PAGE-like Protein in Benign Prostatic

FILE REFERENCE: 44921-5031-WO

CURRENT APPLICATION NUMBER: PCT/US01/23861

PRIOR FILING DATE: 2001-07-31

PRIOR APPLICATION NUMBER: US 60/222,039

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 2

LENGTH: 102

TYPE: PRT

ORGANISM: Homo sapiens

PCT-US01-23861-2

Query Match 62.7%; Score 47; DB 1; Length 102;  
Best Local Similarity 61.5%; Pred. No. 16;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KHKLKOPGNDN 13

Db 90 KHKLKOPGNDN 102

RESULT 14

PCT-US99-09970-2

Sequence 2, Application PC/TUS9909970

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

APPLICANT: McCoy, John M.

APPLICANT: Lavallee, Edward R.

APPLICANT: Collins-Racie, Lisa A.

APPLICANT: Evans, Cheryl

APPLICANT: Merberg, David

APPLICANT: Treacy, Maurice

APPLICANT: Agostino, Michael J.

APPLICANT: Steininger II, Robert J.

APPLICANT: Bowman, Michael R.

APPLICANT: DiBlasio-Smith, Elizabeth

APPLICANT: Widom, Angela

APPLICANT: Genetics Institute, Inc.

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

FILE REFERENCE: GI 6069-74A

CURRENT APPLICATION NUMBER: PCT/US99/09970

PRIOR FILING DATE: 1999-05-07

NUMBER OF SEQ ID NOS: 180

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 2

LENGTH: 102

TYPE: PRT

ORGANISM: Homo sapiens  
PCT-US99-09970-2

Query Match 62.7%; Score 47; DB 1; Length 102;  
Best Local Similarity 61.5%; Pred. No. 16;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKLKOPGDGDP 13  
|||:||||  
Db 90 KHAKTKENGDCGP 102

RESULT 15  
US-09-306-111-2  
; Sequence 2, Application US/09306111  
; GENERAL INFORMATION:

; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John M.  
; APPLICANT: Lavalley, Edward R.  
; APPLICANT: Collins-Racie, Lisa A.  
; APPLICANT: Evans, Cheryl  
; APPLICANT: Merberg, David  
; APPLICANT: Treacy, Maurice  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Steininger II, Robert J.  
; APPLICANT: Bowman, Michael R.  
; APPLICANT: DiBlasio-Smith, Elizabeth  
; APPLICANT: Widom, Angela  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
; FILE REFERENCE: GI 6069-74A  
; CURRENT APPLICATION NUMBER: US/09/306,111  
; CURRENT FILING DATE: 1999-05-06  
; NUMBER OF SEQ ID NOS: 180  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 102  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-306-111-2

Query Match 62.7%; Score 47; DB 17; Length 102;  
Best Local Similarity 61.5%; Pred. No. 16;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKLKOPGDGDP 13  
|||:||||  
Db 90 KHAKTKENGDCGP 102

Search completed: January 29, 2002, 10:56:01  
Job time: 2397 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:08 : Search time 120.95 seconds  
(without alignments)  
3.934 Million cell updates/sec

Title: US-09-763-397A-5  
Perfect score: 75  
Sequence: 1 KHKKLKQPCDGNP 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA-New:\*  
1: /cgn2\_6/ptodata/2/paa/PC7\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	53.3	109	US-09-897-516-4572	Sequence 4572, Ap
2	40	53.3	597	US-09-708-427-11248	Sequence 11248, A
3	40	53.3	628	US-09-708-427-11247	Sequence 11247, A
4	40	53.3	633	US-09-708-427-11246	Sequence 11246, A
5	39	52.0	423	US-09-708-427-6246	Sequence 6246, Ap
6	39	52.0	454	US-09-708-427-6245	Sequence 6245, Ap
7	39	52.0	558	US-09-708-427-6244	Sequence 6244, Ap
8	38	50.7	122	US-09-708-427-43901	Sequence 43901, A
9	38	50.7	125	US-09-708-427-43900	Sequence 43900, A
10	38	50.7	152	US-09-708-427-43899	Sequence 43899, A
11	38	50.7	252	US-10-015-127-13035	Sequence 13035, A
12	38	50.7	287	US-09-708-427-32720	Sequence 32720, A
13	38	50.7	342	US-09-708-427-32719	Sequence 32719, A
14	38	50.7	435	US-09-708-427-32718	Sequence 32718, A
15	38	50.7	856	US-60-337-358-600	Sequence 600, App
16	37	49.3	108	US-09-611-526-4106	Sequence 4106, Ap
17	37	49.3	176	US-09-708-427-15140	Sequence 15140, A
18	37	49.3	193	US-09-708-427-15139	Sequence 15139, A
19	37	49.3	288	US-09-708-427-15138	Sequence 15138, A
20	37	49.3	974	US-09-708-427-16306	Sequence 16306, A
21	37	49.3	993	US-09-708-427-16305	Sequence 16305, A
22	37	49.3	1036	US-09-708-427-16304	Sequence 16304, A
23	36	48.0	135	US-09-620-111B-4689	Sequence 4689, Ap
24	36	48.0	136	US-09-708-427-61433	Sequence 61433, A
25	36	48.0	181	US-09-708-427-49591	Sequence 49591, A
26	36	48.0	195	US-09-760-446A-1794	Sequence 1794, Ap

27	36	48.0	227	5	US-09-969-730-136	Sequence 136, App
28	36	48.0	227	5	US-09-774-639-155	Sequence 155, App
29	36	48.0	252	5	US-09-708-427-59040	Sequence 59040, A
30	36	48.0	261	5	US-09-708-427-59039	Sequence 59039, A
31	36	48.0	291	1	PCT-US01-27760-554	Sequence 554, App
32	36	48.0	291	1	PCT-US01-27760A-554	Sequence 554, App
33	36	48.0	423	5	US-09-708-427-19308	Sequence 19308, A
34	36	48.0	424	5	US-09-708-427-24191	Sequence 24191, A
35	36	48.0	448	5	US-09-708-427-24190	Sequence 24190, A
36	35	46.7	21	5	US-09-821-726-8	Sequence 8, Appl1
37	35	46.7	25	5	US-09-821-726-10	Sequence 10, Appl1
38	35	46.7	36	5	US-09-708-427-74370	Sequence 74370, A
39	35	46.7	42	5	US-09-821-726-6	Sequence 6, Appl1
40	35	46.7	185	5	US-09-989-723-211	Sequence 211, App
41	35	46.7	185	5	US-09-989-724-211	Sequence 211, App
42	35	46.7	185	5	US-09-989-730-211	Sequence 211, App
43	35	46.7	185	5	US-09-990-436-211	Sequence 211, App
44	35	46.7	185	5	US-09-990-444-211	Sequence 211, App
45	35	46.7	185	5	US-09-989-721-211	Sequence 211, App

#### ALIGNMENTS

RESULT 1  
US-09-897-516-4572  
Sequence 4572, Application US/09897516  
GENERAL INFORMATION:  
APPLICANT: Corbin, David R.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Huesling, Joseph E.  
APPLICANT: Krasomil-Osterfeld, Karina C.  
APPLICANT: Malvar, Thomas M.  
APPLICANT: Slater, Steven C.  
APPLICANT: Spiridonov, Sergei  
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51847)B  
CURRENT APPLICATION NUMBER: US/09/897,516  
CURRENT FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/215, 161  
PRIOR FILING DATE: 2000-06-30  
NUMBER OF SEQ ID NOS: 8409  
SEQ ID NO 4572  
LENGTH: 109  
TYPE: PRT  
ORGANISM: Xenorhabdus sp.  
US-09-897-516-4572  
Query Match 53.3%; Score 40; DB 5; Length 109;  
Best Local Similarity 60.0%; Pred. No. 7.4;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KHKKLKQPCD 10  
DB 44 QHKRSKRPED 53  
RESULT 2  
US-09-708-427-11248  
Sequence 11248, Application US/09708427  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
FILE REFERENCE: 2750-1243P  
CURRENT FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 11248  
LENGTH: 597

```
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..597
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..597
; OTHER INFORMATION: Ceres Seq. ID 1822797
US-09-708-427-11248
```

```
Query Match          53.3%; Score 40; DB 5; Length 597;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 KHKRLKQPGDGN 12
      |::|::|::|
Db 56 KNKKLQQRGDTN 67
```

```
RESULT 3
US-09-708-427-11247
; Sequence 11247, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ. ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11247
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..628
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..628
; OTHER INFORMATION: Ceres Seq. ID 1822796
US-09-708-427-11247
```

```
Query Match          53.3%; Score 40; DB 5; Length 628;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 KHKRLKQPGDGN 12
      |::|::|::|
Db 87 KNKKLQQRGDTN 98
```

```
RESULT 4
US-09-708-427-11246
; Sequence 11246, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ. ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11246
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: 1..633
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..633
; OTHER INFORMATION: Ceres Seq. ID 1822795
US-09-708-427-11246
```

```
Query Match          53.3%; Score 40; DB 5; Length 633;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 KHKRLKQPGDGN 12
      |::|::|::|
Db 92 KNKKLQQRGDTN 103
```

```
RESULT 5
US-09-708-427-6246
; Sequence 6246, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ. ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6246
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..425
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..425
; OTHER INFORMATION: Ceres Seq. ID 1815598
US-09-708-427-6246
```

```
Query Match          52.0%; Score 39; DB 5; Length 425;
Best Local Similarity 58.3%; Pred. No. 44;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 KHKRLKQPGDGN 12
      |::|::|::|
Db 94 KKKLKKKEGEGN 105
```

```
RESULT 6
US-09-708-427-6245
; Sequence 6245, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ. ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6245
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..454
; OTHER INFORMATION: Xaa is any amino acid
```



```

: NAME/KEY: misc_feature
: LOCATION: 1..454
: OTHER INFORMATION: Ceres Seq. ID 1815597
US-09-708-427-6245

```

```

Query Match      52.0%; Score 39; DB 5; Length 454;
Best Local Similarity 58.3%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 1 KKKLKQPGDGN 12
    1 1 1 1 1 1 1 1
Db 123 KKKLKQPGDGN 134

```

```

RESULT 7
US-09-708-427-6244
: Sequence 6244, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 6244
: LENGTH: 558
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..558
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..558
: OTHER INFORMATION: Ceres Seq. ID 1815596
US-09-708-427-6244

```

```

Query Match      52.0%; Score 39; DB 5; Length 558;
Best Local Similarity 58.3%; Pred. No. 58;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 1 KKKLKQPGDGN 12
    1 1 1 1 1 1 1 1
Db 227 KKKLKQPGDGN 238

```

```

RESULT 8
US-09-708-427-43901
: Sequence 43901, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 43901
: LENGTH: 122
: TYPE: PRT
: ORGANISM: Zea mays subsp. mays
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..122
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..122
: OTHER INFORMATION: Ceres Seq. ID 1848537

```

```

US-09-708-427-43901

```

```

Query Match      50.7%; Score 38; DB 5; Length 122;
Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 3 KKKLKQPGDGN 11
    1 1 1 1 1 1 1 1
Db 38 KKKLKQPGDGN 46

```

```

RESULT 9
US-09-708-427-43900
: Sequence 43900, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 43900
: LENGTH: 126
: TYPE: PRT
: ORGANISM: Zea mays subsp. mays
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..126
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..126
: OTHER INFORMATION: Ceres Seq. ID 1848536
US-09-708-427-43900

```

```

Query Match      50.7%; Score 38; DB 5; Length 126;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 3 KKKLKQPGDGN 11
    1 1 1 1 1 1 1 1
Db 42 KKKLKQPGDGN 50

```

```

RESULT 10
US-09-708-427-43899
: Sequence 43899, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 43899
: LENGTH: 152
: TYPE: PRT
: ORGANISM: Zea mays subsp. mays
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..152
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..152
: OTHER INFORMATION: Ceres Seq. ID 1848535
US-09-708-427-43899

```

Query Match 50.7%; Score 38; DB 5; Length 152;  
Best Local Similarity 77.8%; Pred. No. 23;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 KKKLKQPGDG 11  
|:|:|:|:|:|  
Db 68 KRLKOSGDG 76

RESULT 11  
US-10-015-127-13035  
; Sequence 13035, Application US/10015127  
; GENERAL INFORMATION:  
; APPLICANT: Bower, Stanley G.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Sphingomonas elodea genome sequences and uses thereof  
; FILE REFERENCE: 38-10(15806)B  
; CURRENT APPLICATION NUMBER: US/10/015,127  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: US 60/252,455  
; PRIOR FILING DATE: 2000-11-22  
; NUMBER OF SEQ ID NOS: 14357  
; SEQ ID NO 13035  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Sphingomonas elodea  
US-10-015-127-13035

Query Match 50.7%; Score 38; DB 6; Length 252;  
Best Local Similarity 45.3%; Pred. No. 38;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 KKKLKQPGDG 11  
|:|:|:|:|:|  
Db 241 QHRKIAERGG 251

RESULT 12  
US-09-708-427-32720  
; Sequence 32720, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: Patentln version 3.1  
; SEQ ID NO 32720  
; LENGTH: 287  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..287  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..287  
; OTHER INFORMATION: Ceres Seq. ID 1834864  
US-09-708-427-32720

Query Match 50.7%; Score 38; DB 5; Length 287;  
Best Local Similarity 58.3%; Pred. No. 44;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KKKLKQPGDGN 12  
|:|:|:|:|:|  
Db 46 KVEVEQEGDGN 57

RESULT 13  
US-09-708-427-32719  
; Sequence 32719, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: Patentln version 3.1  
; SEQ ID NO 32719  
; LENGTH: 342  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..342  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..342  
; OTHER INFORMATION: Ceres Seq. ID 1834863  
US-09-708-427-32719

Query Match 50.7%; Score 38; DB 5; Length 342;  
Best Local Similarity 58.3%; Pred. No. 52;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KKKLKQPGDGN 12  
|:|:|:|:|:|  
Db 101 KVEVEQEGDGN 112

RESULT 14  
US-09-708-427-32718  
; Sequence 32718, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: Patentln version 3.1  
; SEQ ID NO 32718  
; LENGTH: 435  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..435  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..435  
; OTHER INFORMATION: Ceres Seq. ID 1834862  
US-09-708-427-32718

Query Match 50.7%; Score 38; DB 5; Length 435;  
Best Local Similarity 58.3%; Pred. No. 67;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KKKLKQPGDGN 12  
|:|:|:|:|:|  
Db 194 KVEVEQEGDGN 205

RESULT 15  
US-60-337-358-600

```

: Sequence 600, Application US/60337358
: GENERAL INFORMATION:
: APPLICANT: Agarwal, Ameeta K.
: APPLICANT: Ahrens, Jeffrey E.
: APPLICANT: Ball, James A.
: APPLICANT: Banu, G.
: APPLICANT: Bell, Erin
: APPLICANT: Boddupalli, Raghava
: APPLICANT: Chomet, Paul S.
: APPLICANT: Daly, Mackenzie
: APPLICANT: Deikman, Jill
: APPLICANT: Deng, Mollan
: APPLICANT: Dong, Jinzhuo
: APPLICANT: Duff, Stephen M.
: APPLICANT: Edgerton, Michael D.
: APPLICANT: Galligan, Meghan M.
: APPLICANT: Hinchey, Brenda S.
: APPLICANT: Huang, Shishleh
: APPLICANT: Johnson, Richard G.
: APPLICANT: Jung, Vincent
: APPLICANT: Kretzmer, Keith A.
: APPLICANT: Laccetti, Lucille B.
: APPLICANT: Lal, Chao-Qiang
: APPLICANT: Lee, Gary
: APPLICANT: Lin, Jie-Yi
: APPLICANT: Liu, Jingdong
: APPLICANT: Lu, Bin
: APPLICANT: Luechy, Michael M.
: APPLICANT: Lund, Adrian
: APPLICANT: Madson, Linda L.
: APPLICANT: Malloy, Kathleen A.
: APPLICANT: McKiel, Christine L.
: APPLICANT: Miller, Philip W.
: APPLICANT: Padmathi, Machikanti
: APPLICANT: Patnell, Laurence D.
: APPLICANT: Sheridan, Paul
: APPLICANT: Sherman, Paul L.
: APPLICANT: Start, William G.
: APPLICANT: Teensesen, Dan
: APPLICANT: Vidya, K.R.
: APPLICANT: Wang, Haiyun
: APPLICANT: Xin, Zhanguo
: APPLICANT: Xu, Nanfei
: APPLICANT: Yang, Chunzhi
: APPLICANT: Zeng, Xiaoping
: APPLICANT: Zhang, Qiang
: APPLICANT: Zhao, Yajuan
: APPLICANT: Zhou, Li
: TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
: FILE REFERENCE: 38-15(52796)A
: CURRENT APPLICATION NUMBER: US/60/337,358
: CURRENT FILING DATE: 2001-12-04
: NUMBER OF SEQ ID NOS: 745
: SEQ ID NO 600
: LENGTH: 856
: TYPE: PRT
: ORGANISM: Zea mays
: US-60-337-358-600

```

```

Query Match          50.7%: Score 38: DB 7: Length 856:
Best Local Similarity 61.5%: Pred. No. 1.3e+02;
Matches 8: Conservative 2; Mismatches 1; Indels 2; Gaps 1;
Oy 1 KKKLKPQ-DG 11
   ||| |::|| |
Db 539 KKKLREPQPDG 551

```

**THIS PAGE BLANK (USPTO)**

GenCore version: 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: January 29, 2002, 10:55:59 ; Search time 1760.55 Seconds  
(without alignments)  
1.893 Million cell updates/sec

Title: US-09-763-397A-4  
Perfect score: 69  
Sequence: 1 NANPNANPNANP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Pending\_Patents\_AA\_Main: \*  
1: /cgn2\_6/ptodata/2/paa/PCRTUS\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep: \*  
6: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep: \*  
7: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep: \*  
8: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep: \*  
9: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep: \*  
10: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep: \*  
11: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
12: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep: \*  
13: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep: \*  
14: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep: \*  
15: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep: \*  
16: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep: \*  
17: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep: \*  
18: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep: \*  
19: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep: \*  
20: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep: \*  
21: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep: \*  
22: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep: \*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep: \*  
24: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69	100.0	12	1	PCT-US99-03055-105
2	69	100.0	12	3	Sequence 105, App
3	69	100.0	12	1	Sequence 9, Appl1
4	69	100.0	12	7	Sequence 11, Appl1
5	69	100.0	12	7	Sequence 4, Appl1
6	69	100.0	12	11	Sequence 8, Appl1
7	69	100.0	12	13	Sequence 1, Appl1
8	69	100.0	12	13	Sequence 1, Appl1
9	69	100.0	12	14	Sequence 34, Appl1

10	69	100.0	12	17	US-09-300-959-39	Sequence 39, Appl1
11	69	100.0	12	21	US-09-763-397A-4	Sequence 4, Appl1
12	69	100.0	13	3	US-07-958-321-13	Sequence 13, Appl1
13	69	100.0	13	5	US-08-138-514-8	Sequence 8, Appl1
14	69	100.0	13	7	US-08-387-932-3	Sequence 3, Appl1
15	69	100.0	13	11	US-08-773-675-11	Sequence 11, Appl1
16	69	100.0	13	11	US-08-773-675A-10	Sequence 10, Appl1
17	69	100.0	13	11	US-08-773-675B-10	Sequence 10, Appl1
18	69	100.0	13	17	US-09-389-482-10	Sequence 10, Appl1
19	69	100.0	16	3	US-07-958-321-12	Sequence 12, Appl1
20	69	100.0	16	8	US-08-466-755-6	Sequence 6, Appl1
21	69	100.0	16	11	US-08-773-675-10	Sequence 10, Appl1
22	69	100.0	16	11	US-08-773-675A-9	Sequence 9, Appl1
23	69	100.0	16	11	US-08-773-675B-9	Sequence 9, Appl1
24	69	100.0	16	17	US-09-310-462-18	Sequence 18, Appl1
25	69	100.0	16	17	US-09-389-482-9	Sequence 9, Appl1
26	69	100.0	16	17	US-09-454-204A-26	Sequence 26, Appl1
27	69	100.0	16	18	US-08-944-147-19	Sequence 19, Appl1
28	69	100.0	20	13	US-08-998-335-4	Sequence 4, Appl1
29	69	100.0	28	13	US-08-998-335-4	Sequence 4, Appl1
30	69	100.0	28	13	US-08-998-335-4	Sequence 4, Appl1
31	69	100.0	28	14	US-09-060-450-10	Sequence 10, Appl1
32	69	100.0	36	1	PCT-US99-139758-104	Sequence 104, App
33	69	100.0	36	1	PCT-US99-139758-105	Sequence 104, App
34	69	100.0	36	21	US-09-701-588-105	Sequence 105, App
35	69	100.0	36	21	US-09-701-588-105	Sequence 105, App
36	69	100.0	48	14	US-09-060-450-11	Sequence 11, Appl1
37	69	100.0	68	4	US-08-029-666A-12	Sequence 12, Appl1
38	69	100.0	68	13	US-08-982-965-13	Sequence 13, Appl1
39	69	100.0	130	8	US-08-466-755-7	Sequence 7, Appl1
40	69	100.0	163	3	US-07-941-654-2	Sequence 2, Appl1
41	69	100.0	163	7	US-08-397-128-2	Sequence 2, Appl1
42	69	100.0	164	3	US-07-941-654-1	Sequence 1, Appl1
43	69	100.0	164	7	US-08-397-128-1	Sequence 1, Appl1
44	69	100.0	164	22	US-09-826-513-1	Sequence 1, Appl1
45	69	100.0	229	18	US-09-454-204A-41	Sequence 41, Appl1

## ALIGNMENTS

RESULT 1  
PCT-US99-03055-105  
Sequence 105, Application PC/TUS9903055  
GENERAL INFORMATION:  
APPLICANT: Birkett, Ashley J.  
TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and  
FILE REFERENCE: SYN-101 4564/69529  
CURRENT APPLICATION NUMBER: PCT/US99/03055  
EARLIER FILING DATE: 1999-02-11  
EARLIER APPLICATION NUMBER: 60/074537  
NUMBER OF SEQ ID NOS: 113  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 105  
LENGTH: 12  
TYPE: PPT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Cytochrome  
PUBLICATION INFORMATION:  
JOURNAL: Science  
VOLUME: 228  
PAGES: 1436-1440  
DATE: 1985  
PCT-US99-03055-105

Query Match 100.0%; Score 69; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNNANP 12  
Db 1 NANPNNANP 12

RESULT 2  
US-07-958-321-9  
; Sequence 9, Application US/07958321  
; GENERAL INFORMATION:  
; APPLICANT: Perham, Richard N.  
; APPLICANT: Willis, Anne E.  
; APPLICANT: Greenwood, Judith  
; TITLE OF INVENTION: ENGINEERED BACTERIOPHAGES AND VACCINES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: POMS, SMITH, LANDE & ROSE  
; STREET: 2121 Avenue of the Stars  
; CITY: Los Angeles  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 90067  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/958,321  
; FILING DATE: 19921222  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9022190.4  
; FILING DATE: 12-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB91/01785  
; FILING DATE: 14-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maclean, Kurt A.  
; REGISTRATION NUMBER: 31,118  
; REFERENCE/DOCKET NUMBER: 89-208  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 714-263-8250  
; TELEFAX: 714-263-8250  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-958-321-9

Query Match 100.0%; Score 69; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNNANP 12  
Db 1 NANPNNANP 12

RESULT 3  
US-07-958-321-11  
; Sequence 11, Application US/07958321  
; GENERAL INFORMATION:  
; APPLICANT: Perham, Richard N.  
; APPLICANT: Willis, Anne E.  
; APPLICANT: Greenwood, Judith  
; TITLE OF INVENTION: ENGINEERED BACTERIOPHAGES AND VACCINES  
; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: POMS, SMITH, LANDE & ROSE  
; STREET: 2121 Avenue of the Stars  
; CITY: Los Angeles  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 90067  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/958,321  
; FILING DATE: 19921222  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9022190.4  
; FILING DATE: 12-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB91/01785  
; FILING DATE: 14-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maclean, Kurt A.  
; REGISTRATION NUMBER: 31,118  
; REFERENCE/DOCKET NUMBER: 89-208  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 714-263-8250  
; TELEFAX: 714-263-8250  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-958-321-11

Query Match 100.0%; Score 69; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNNANP 12  
Db 1 NANPNNANP 12

RESULT 4  
US-08-387-932-4  
; Sequence 4, Application US/08387932  
; GENERAL INFORMATION:  
; APPLICANT: COMIS, ALFIO  
; APPLICANT: FISCHER, PETER  
; APPLICANT: TYLER, MARGARET I  
; TITLE OF INVENTION: Retro-, Inverso-, and Retro-Inverso  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr., P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: United States of America  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Vers. #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/387,932  
; FILING DATE:  
; CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU93/00441  
FILING DATE: 27-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU 4374  
FILING DATE: 27-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: CHC2USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium falciparum  
DEVELOPMENTAL STAGE: SPOROZOITE  
PUBLICATION INFORMATION:  
DOCUMENT NUMBER: AU P14374  
FILING DATE: 27-AUG-1992  
US-08-387-932-4

Query Match 100.0%; Score 69; DB 7; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
|||||  
DB 1 NANPNANPNANP 12

RESULT 5  
US-08-773-675-8  
Sequence 8, Application US/08773675  
GENERAL INFORMATION:  
APPLICANT: PERHAM, RICHARD N.  
APPLICANT: WILTS, ANNE  
APPLICANT: GREENWOOD, JUDITH  
TITLE OF INVENTION: ENGINEERED BACTERIOPHAGES AND VACCINES  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHAYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/773,675  
FILING DATE: 24-DEC-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/604,958  
FILING DATE: 03-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MITCHARD, LEONARD C.  
REGISTRATION NUMBER: 29,009

REFERENCE/DOCKET NUMBER: 179-13  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-773-675-8

Query Match 100.0%; Score 69; DB 11; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
|||||  
DB 1 NANPNANPNANP 12

RESULT 6  
US-08-998-335-1  
Sequence 1, Application US/08998335  
GENERAL INFORMATION:  
APPLICANT: Nardin, Elizabeth  
APPLICANT: Nussenzweig, Ruth S.  
APPLICANT: Rose, Keith  
TITLE OF INVENTION: POLYOXIME-BASED ANTI-MALARIAL VACCINES  
FILE REFERENCE: 5986/1C599US1  
CURRENT APPLICATION NUMBER: US/08/998,335  
CURRENT FILING DATE: 1997-12-24  
EARLIER APPLICATION NUMBER: US 60/034,506  
EARLIER FILING DATE: 1996-12-24  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 12  
TYPE: PRT  
ORGANISM: P. falciparum  
US-08-998-335-1

Query Match 100.0%; Score 69; DB 13; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
|||||  
DB 1 NANPNANPNANP 12

RESULT 7  
US-08-998-335-1  
Sequence 1, Application US/08998335A  
GENERAL INFORMATION:  
APPLICANT: Nardin, Elizabeth  
APPLICANT: Nussenzweig, Ruth S.  
APPLICANT: Rose, Keith  
TITLE OF INVENTION: POLYOXIME-BASED ANTI-MALARIAL VACCINES  
FILE REFERENCE: 5986/1C599-US1  
CURRENT APPLICATION NUMBER: US/08/998,335A  
CURRENT FILING DATE: 1997-12-24  
EARLIER APPLICATION NUMBER: 60/034,506  
EARLIER FILING DATE: 1996-12-24  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 12  
TYPE: PRT  
ORGANISM: P. falciparum  
US-08-998-335-1

Query Match 100.0%; Score 69; DB 13; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 1 NANPNANPNANP 12

RESULT 8  
US-09-060-450-1  
; Sequence 1, Application US/09060450  
; GENERAL INFORMATION:  
; APPLICANT: Nardin, Elizabeth  
; TITLE OF INVENTION: UNIVERSAL T-CELL EPITOPES FOR ANTI-MALARIAL VACCINES  
; FILE REFERENCE: 5986/1B615-0S1  
; CURRENT APPLICATION NUMBER: US/09/060.450  
; CURRENT FILING DATE: 1998-01-21  
; EARLIER APPLICATION NUMBER: 60/033.916  
; EARLIER FILING DATE: 1997-01-21  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: P. falciparum  
US-09-060-450-1

Query Match 100.0%; Score 69; DB 14; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 1 NANPNANPNANP 12

RESULT 9  
US-09-074-410-34  
; Sequence 34, Application US/09074410  
; GENERAL INFORMATION:  
; APPLICANT: AGRAWAL, Babita  
; APPLICANT: KRANTZ, Mark J.  
; APPLICANT: REDDISH, Mark A.  
; APPLICANT: LONGENECKER, B. Michael  
; TITLE OF INVENTION: METHOD FOR GENERATING ACTIVATED T-CELLS  
; TITLE OF INVENTION: AND ANTIGEN-PULSED ANTIGEN-PRESENTING CELLS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY & LARDNER  
; STREET: 3000 K Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/074.410  
; FILING DATE: 08-MAY-1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/045.949  
; FILING DATE: 08-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saxe, Bernhard D.

REGISTRATION NUMBER: 28.665  
REFERENCE/DOCKET NUMBER: 042881/0114  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-074-410-34

Query Match 100.0%; Score 69; DB 14; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 1 NANPNANPNANP 12

RESULT 10  
US-09-300-959-39  
; Sequence 39, Application US/09300959  
; GENERAL INFORMATION:  
; APPLICANT: Zanetti, Maurizio  
; TITLE OF INVENTION: Somatic Transgene Immunization and Related Methods  
; FILE REFERENCE: P-2A 3519  
; CURRENT APPLICATION NUMBER: US/09/300.959  
; CURRENT FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: US 60/083.154  
; PRIOR FILING DATE: 1998-04-27  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-300-959-39

Query Match 100.0%; Score 69; DB 17; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NANPNANPNANP 12  
Db 1 NANPNANPNANP 12

RESULT 11  
US-09-763-397A-4  
; Sequence 4, Application US/09763397A  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services, Centers for  
; APPLICANT: Control and Prevention  
; APPLICANT: Lal, Altaf A.  
; APPLICANT: Ping Shi, Ya  
; APPLICANT: Hasnain, Seved E.  
; TITLE OF INVENTION: Recombinant Multivalent Malarial Vaccine Against Plasmodium Fa  
; FILE REFERENCE: 6395-57049  
; CURRENT APPLICATION NUMBER: US/09/763.397A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/097.703  
; PRIOR FILING DATE: 1998-08-21  
; PRIOR APPLICATION NUMBER: PCT / US99/18869



PRIOR FILING DATE: 1999-08-19  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Plasmodium falciparum  
US-09-763-397A-4

Query Match 100.0%; Score 69; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NANPNANPNANP 12  
Db 1 NANPNANPNANP 12

RESULT 12

US-07-958-321-13  
Sequence 13, Application US/07958321

GENERAL INFORMATION:  
APPLICANT: Pertham, Richard N.  
APPLICANT: Willis, Anne E.  
TITLE OF INVENTION: ENGINEERED BACTERIOPHAGES AND VACCINES  
TITLE OF INVENTION: CONTAINING THEM  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: POMS, SMITH, LANDE & ROSE  
STREET: 2121 Avenue of the Stars  
CITY: Los Angeles  
STATE: CA  
COUNTRY: USA  
ZIP: 90067

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/958,321  
FILING DATE: 19921222  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9022190.4  
FILING DATE: 12-OCT-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB91/01785  
FILING DATE: 14-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MacLean, Kurt A.  
REGISTRATION NUMBER: 31,118  
REFERENCE/DOCKET NUMBER: 89-208  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 714-263-8250  
TELEFAX: 714-263-8260  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: P. falciparum  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..13  
OTHER INFORMATION: /note="Chemically synthesized  
OTHER INFORMATION: peptide used to check reactivity of antibodies  
OTHER INFORMATION: raised against fdm1."

US-07-958-321-13

Query Match 100.0%; Score 69; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NANPNANPNANP 12  
Db 1 NANPNANPNANP 12

RESULT 13

US-08-138-514-8  
Sequence 8, Application US/08138514

GENERAL INFORMATION:  
APPLICANT: Cunningham, Barry  
APPLICANT: Hannah, John  
APPLICANT: Tolman, Richard L  
TITLE OF INVENTION: ANNUAL ANTIGEN SCAFFOLDS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: United States  
ZIP: 07065

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/138,514  
FILING DATE: 15-OCT-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Bengten, Gerard H.  
REGISTRATION NUMBER: 35,745  
REFERENCE/DOCKET NUMBER: 19008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-3901  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
FEATURE:  
NAME/KEY: Cross-links  
LOCATION: 2..4  
NAME/KEY: Cross-links  
LOCATION: 6..8  
NAME/KEY: Cross-links  
LOCATION: 10..12  
NAME/KEY: Modified-site  
LOCATION: 13  
OTHER INFORMATION: /label= amide  
US-08-138-514-8

Query Match 100.0%; Score 69; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
| | | | |  
Db 2 NANPNANPNANP 13

RESULT 14  
US-08-387-932-3  
; Sequence 3, Application US/08387932  
; GENERAL INFORMATION:  
; APPLICANT: COMIS, ALFEO  
; APPLICANT: FISCHER, PETER  
; APPLICANT: TYLER, MARGARET I  
; TITLE OF INVENTION: Retro-, Inverso-, and Retro-Inverso  
; TITLE OF INVENTION: Synthetic Peptide Analogues  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntrl., P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: United States of America  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Vers. #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/387,932  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/AU93/00441  
; FILING DATE: 27-AUG-1993  
; PRIOR APPLICATION DATA: AU 4374  
; FILING DATE: 27-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: GNC20USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; TEXT:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: Internal  
; ORIGINAL SOURCE:  
; ORGANISM: Plasmodium falciparum  
; DEVELOPMENTAL STAGE: SPOROZOITE  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1..2  
; OTHER INFORMATION: /label-A  
; OTHER INFORMATION: /note="RESIDUE 1 IS AN EXTRA CYSTEINE  
; OTHER INFORMATION: RESIDUE ADDED TO THE N-TERMINUS OF THE  
; OTHER INFORMATION: PEPTIDE"  
; PUBLICATION INFORMATION:  
; DOCUMENT NUMBER: AU PL4374  
; FILING DATE: 27-AUG-1992  
; US-08-387-932-3

Query Match 100.0%; Score 69; DB 7; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NANPNANPNANP 12  
| | | | |  
Db 2 NANPNANPNANP 13

RESULT 15  
US-08-773-675-11  
; Sequence 11, Application US/08773675  
; GENERAL INFORMATION:  
; APPLICANT: PERHAM, RICHARD N.  
; APPLICANT: WILITS, ANNE  
; APPLICANT: GREENWOOD, JUDITH  
; TITLE OF INVENTION: ENGINEERED BACTERIOPHAGES AND VACCINES  
; TITLE OF INVENTION: CONTAINING THEM  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: NIXON & VANDERHAYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/773,675  
; FILING DATE: 24-DEC-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/604,958  
; FILING DATE: 03-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MITCHARD, LEONARD C.  
; REGISTRATION NUMBER: 29,009  
; REFERENCE/DOCKET NUMBER: 179-13  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-773-675-11

Query Match 100.0%; Score 69; DB 11; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: January 29, 2002, 10:56:00  
Job time: 2396 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2002, 10:58:07 ; Search time 120.95 Seconds  
(without alignments)  
3.631 Million cell updates/sec

Title: US-09-763-397A-4  
Perfect score: 69  
Sequence: 1 NANPNANPNANP 12

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 173191 seqs, 36597120 residues

Total number of hits satisfying chosen parameters: 173191

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	100.0	68	5	US-09-938-406-5
2	54	78.3	28	5	US-09-543-696B-23
3	49	71.0	818	5	US-09-897-516-5117
4	49	71.0	1477	5	US-09-482-500-1
5	49	71.0	1477	5	US-09-482-500A-1
6	49	71.0	1704	5	US-09-543-696B-6
7	47	68.1	367	5	US-09-971-118-2
8	46	66.7	757	5	US-09-708-427-26420
9	46	66.7	953	5	US-09-708-427-26419
10	45	65.2	424	5	US-09-708-427-78695
11	45	65.2	521	5	US-09-708-427-78694
12	44	63.8	346	5	US-09-708-427-31468
13	43	62.3	342	5	US-09-708-427-3419
14	43	62.3	343	5	US-09-708-427-3418
15	43	62.3	356	5	US-09-708-427-3417
16	42	60.9	263	5	US-09-708-427-34465
17	41	59.4	264	6	US-10-015-127-10155
18	41	59.4	378	5	US-09-708-427-2211
19	41	59.4	391	6	US-10-015-127-12005
20	41	59.4	425	5	US-09-708-427-2210
21	41	59.4	442	5	US-09-708-427-2209
22	41	59.4	487	5	US-09-926-805-6
23	40	58.0	131	5	US-09-708-427-37918
24	40	58.0	146	5	US-09-708-427-37916
25	40	58.0	280	5	US-09-708-427-14917
26	40	58.0	288	5	US-09-708-427-60910

27	39.5	57.2	236	5	US-09-708-427-53326	Sequence 53326, A
28	39.5	57.2	303	5	US-09-708-427-53325	Sequence 53325, A
29	39	56.5	186	5	US-09-760-446A-2229	Sequence 2229, Ap
30	39	56.5	207	5	US-09-760-446A-1187	Sequence 1187, Ap
31	39	56.5	309	5	US-09-760-446A-1507	Sequence 1507, Ap
32	39	56.5	388	5	US-09-646-673A-104	Sequence 104, App
33	39	56.5	402	5	US-09-708-427-1204	Sequence 1204, Ap
34	39	56.5	406	5	US-09-708-427-4718	Sequence 4718, Ap
35	39	56.5	431	5	US-09-981-353-7	Sequence 7, Appl1
36	39	56.5	432	5	US-09-760-446A-1902	Sequence 1902, Ap
37	39	56.5	451	5	US-09-708-427-1896	Sequence 1896, Ap
38	39	56.5	466	5	US-09-708-427-4717	Sequence 4717, Ap
39	39	56.5	536	5	US-09-708-427-1895	Sequence 1895, Ap
40	39	56.5	551	5	US-09-708-427-4716	Sequence 4716, Ap
41	39	56.5	577	5	US-09-708-427-1894	Sequence 1894, Ap
42	39	56.5	1333	5	US-09-708-427-19136	Sequence 19136, A
43	38.5	55.8	446	5	US-09-708-427-2109	Sequence 2109, Ap
44	38.5	55.8	481	5	US-09-708-427-2108	Sequence 2108, Ap
45	38.5	55.8	499	5	US-09-708-427-2107	Sequence 2107, Ap

ALIGNMENTS

RESULT 1

US-09-938-406-5  
; Sequence 5, Application US/09938406  
; GENERAL INFORMATION:  
; APPLICANT: Lowell, George  
; APPLICANT: Vancott, Thomas  
; APPLICANT: Birx, Deborah  
; TITLE OF INVENTION: PROTEIN AND PEPTIDE VACCINES FOR INDUCING MUCOSAL IMMUNITY  
; FILE REFERENCE: 40646-20002.10  
; CURRENT APPLICATION NUMBER: US/09/938,406  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 09/214,701  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: PCT/US 97/12253  
; PRIOR FILING DATE: 1997-07-10  
; PRIOR APPLICATION NUMBER: US 60/021,687  
; PRIOR FILING DATE: 1996-07-10  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: P. falciparum  
US-09-938-406-5

Query Match 100.0%; Score 69; DB 5; Length 68;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NANPNANPNANP 12

Db 1 NANPNANPNANP 12

RESULT 2

US-09-543-696B-23  
; Sequence 23, Application US/09543696B  
; GENERAL INFORMATION:  
; APPLICANT: Potempa, Jan S.  
; APPLICANT: Travis, James  
; APPLICANT: Genco, Caroline A.  
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING PORPHYROMONAS GINGIVALIS PEPTIDES AND METHODS  
; FILE REFERENCE: 235.00120120  
; CURRENT APPLICATION NUMBER: US/09/543,696B  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: 08/822,324

; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: 60/013,945  
; PRIOR FILING DATE: 1996-03-22  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Peptide  
US-09-543-696B-23

Query Match 78.3%; Score 54; DB 5; Length 28;  
Best Local Similarity 75.0%; Pred. No. 0.012;  
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NANPNANPNP 12  
Db 10 NPENPNPNP 21  
| | | | | | | |

## RESULT 3

US-09-897-516-5117  
; Sequence 5117, Application US/09897516  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Huesling, Joseph E.  
; APPLICANT: Krasomil-Osterfeld, Karina C.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)B  
; CURRENT APPLICATION NUMBER: US/09/897,516  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/215, 161  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8409  
; SEQ ID NO 5117  
; LENGTH: 818  
; TYPE: PRT  
; ORGANISM: Xenorhabdus sp.  
US-09-897-516-5117

Query Match 71.0%; Score 49; DB 5; Length 818;  
Best Local Similarity 80.0%; Pred. No. 2.4;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NPANPNANP 12  
Db 681 NPENPNPNP 690  
| | | | | | | |

## RESULT 4

US-09-482-500-1  
; Sequence 1, Application US/09482500  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Imamura, Takahisa  
; APPLICANT: Potempa, Jan  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BLOOD COAGULATION  
; FILE REFERENCE: 235.00160101  
; CURRENT APPLICATION NUMBER: US/09/482,500  
; CURRENT FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: US 60/115,869  
; PRIOR FILING DATE: 1999-01-13  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1  
; LENGTH: 1477  
; TYPE: PRT  
; ORGANISM: Porphyromonas gingivalis  
US-09-482-500-1

Query Match 71.0%; Score 49; DB 5; Length 1477;  
Best Local Similarity 66.7%; Pred. No. 4.4;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NANPNANPNP 12  
Db 720 NGTPNPNPNP 731  
| | | | | | | |

## RESULT 5

US-09-482-500A-1  
; Sequence 1, Application US/09482500A  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Imamura, Takahisa  
; APPLICANT: Potempa, Jan  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BLOOD COAGULATION  
; FILE REFERENCE: 235.00160101  
; CURRENT APPLICATION NUMBER: US/09/482,500A  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/115,869  
; PRIOR FILING DATE: 1999-01-13  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 1477  
; TYPE: PRT  
; ORGANISM: Porphyromonas gingivalis  
US-09-482-500A-1

Query Match 71.0%; Score 49; DB 5; Length 1477;  
Best Local Similarity 66.7%; Pred. No. 4.4;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NANPNANPNP 12  
Db 720 NGTPNPNPNP 731  
| | | | | | | |

## RESULT 6

US-09-543-696B-6  
; Sequence 6, Application US/09543696B  
; GENERAL INFORMATION:  
; APPLICANT: Potempa, Jan S.  
; APPLICANT: Travis, James  
; APPLICANT: Genco, Caroline A.  
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING PORPHYROMONAS  
; FILE REFERENCE: 235.00120120  
; CURRENT APPLICATION NUMBER: US/09/543,696B  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: 08/822,324  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: 60/013,945  
; PRIOR FILING DATE: 1996-03-22  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 1704  
; TYPE: PRT  
; ORGANISM: Porphyromonas gingivalis  
US-09-543-696B-6

Query Match 71.0%; Score 49; DB 5; Length 1704;

Best Local Similarity 66.7%; Pred. No. 5.1;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NANPNANPNP 12  
| | | | | | |  
Db 947 NGTPNPNPNP 958

RESULT 7

US-09-971-118-2  
; Sequence 2, Application: US/09971118  
; GENERAL INFORMATION:  
; APPLICANT: DELANEY, ALLEN  
; APPLICANT: YOGANATHAN, THILLAINATHAN  
; TITLE OF INVENTION: SGK2 AND ITS USES  
; FILE REFERENCE: KINF025CIP  
; CURRENT APPLICATION NUMBER: US/09/971,118  
; CURRENT FILING DATE: 2001-10-03  
; PRIOR APPLICATION NUMBER: PCT/US01/21479  
; PRIOR FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: 60/237,419  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 367  
; TYPE: PRT  
; ORGANISM: Homp sapiens  
US-09-971-118-2

Query Match 68.1%; Score 47; DB 5; Length 367;  
Best Local Similarity 66.7%; Pred. No. 2.1;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NANPNANPNP 12  
| | | | | | |  
Db 21 NLGPSANPNAQ 32

RESULT 8

US-09-708-427-26420  
; Sequence 26420, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26420  
; LENGTH: 757  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..757  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..757  
; OTHER INFORMATION: Ceres Seq. ID 1818048  
US-09-708-427-26420

Query Match 66.7%; Score 46; DB 5; Length 757;  
Best Local Similarity 66.7%; Pred. No. 6.3;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NANPNANPNP 12  
| | | | | | |  
Db 6 NLNYPNPNQNP 17

RESULT 9

US-09-708-427-26419  
; Sequence 26419, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26419  
; LENGTH: 953  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..953  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..953  
; OTHER INFORMATION: Ceres Seq. ID 1818047  
US-09-708-427-26419

Query Match 66.7%; Score 46; DB 5; Length 953;  
Best Local Similarity 66.7%; Pred. No. 7.9;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NANPNANPNP 12  
| | | | | | |  
Db 202 NLNYPNPNQNP 213

RESULT 10

US-09-708-427-78695  
; Sequence 78695, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 78695  
; LENGTH: 424  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..424  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..424  
; OTHER INFORMATION: Ceres Seq. ID 1963464  
US-09-708-427-78695

Query Match 65.2%; Score 45; DB 5; Length 424;  
Best Local Similarity 80.0%; Pred. No. 4.9;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NANPNANPNA 10  
| | | | | | |  
Db 106 NANPNANSNS 115

RESULT 11

US-09-708-427-78694  
; Sequence 78694, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 78694  
; LENGTH: 521  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..521  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..521  
; OTHER INFORMATION: Ceres Seq. ID 1963463  
US-09-708-427-78694

Query Match 65.2%; Score 45; DB 5; Length 521;  
Best Local Similarity 80.0%; Pred. No. 6;  
Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 1 NANNANPNA 10  
|||||||I:  
Db 203 NANNPANSNS 212

RESULT 12  
US-09-708-427-31468  
; Sequence 31468, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31468  
; LENGTH: 346  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..346  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..346  
; OTHER INFORMATION: Ceres Seq. ID 1832153  
US-09-708-427-31468

Query Match 63.8%; Score 44; DB 5; Length 346;  
Best Local Similarity 70.0%; Pred. No. 5.6;  
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 0;

QY 3 NPANPNANP 12  
:|||I||I  
Db 5 DPNPNPNP 14

RESULT 13  
US-09-708-427-3419  
; Sequence 3419, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3419  
; LENGTH: 342  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..342  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..342  
; OTHER INFORMATION: Ceres Seq. ID 1810474  
US-09-708-427-3419

Query Match 62.3%; Score 43; DB 5; Length 342;  
Best Local Similarity 34.6%; Pred. No. 7.8;  
Matches 9; Conservative 3; Mismatches 0; Indels 14; Gaps 1;

QY 1 NANNPN-----ANPNANP 12  
:|||||:  
Db 83 NSNPPESSSNPPDSSSNPN 108

RESULT 14  
US-09-708-427-3418  
; Sequence 3418, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3418  
; LENGTH: 343  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..343  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..343  
; OTHER INFORMATION: Ceres Seq. ID 1810473  
US-09-708-427-3418

Query Match 62.3%; Score 43; DB 5; Length 343;  
Best Local Similarity 34.6%; Pred. No. 7.9;  
Matches 9; Conservative 3; Mismatches 0; Indels 14; Gaps 1;

QY 1 NANNPN-----ANPNANP 12  
:|||||:  
Db 84 NSNPPESSSNPPDSSSNPN 109

RESULT 15  
US-09-708-427-3417  
; Sequence 3417, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3417  
; LENGTH: 342  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..342  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..342  
; OTHER INFORMATION: Ceres Seq. ID 1810474  
US-09-708-427-3417



FILE REFERENCE: 2750-1243P  
CURRENT APPLICATION NUMBER: US/09/708.427  
CURRENT FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3417  
LENGTH: 356  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..356  
OTHER INFORMATION: Xaa is any amino acid  
NAME/KEY: misc\_feature  
LOCATION: 1..356  
OTHER INFORMATION: Ceres Seq. ID 1810472  
US-09-708-427-3417

Query Match 62.3%; Score 43; DB 5; Length 356;  
Best Local Similarity 34.6%; Pred. NO. 8.2;  
Matches 9; Conservative 3; Mismatches 0; Indels 14; Gaps 1;

Oy 1 NANPN-----ANPNAP 12  
I:III  
:III:II  
Db 97 NSNPNPSSSNPNPPDSSSNPNSNP 122

Search completed: January 29, 2002, 10:58:08  
Job time: 2504 sec

**THIS PAGE BLANK (USPTO)**